Fitting Feelings into Frameworks:
An examination of the involvement of
primary stakeholders
in the design and use of outcome indicators
and evaluation use
for development interventions

A thesis presented in partial fulfilment of the requirements for the degree
of Master of International Development

Massey University
Palmerston North, New Zealand

Shupayi Mpunga

2015

Abstract

The aim of this thesis is to examine the involvement of primary stakeholders in the design and use of outcome indicators, and evaluation use for development interventions. It seeks to understand the characteristics of indicators developed by or with primary stakeholders, the dynamics of such involvement, and what is achieved in such practice.

The experiences and achievements recorded in eight international case examples and one local case study provide insights into the practical considerations for using participatory approaches to design outcome indicators. Although there is a traceable body of literature that provides theoretical guidance on participatory indicator design, there is very little in the way of instruction based on practice.

Primary stakeholder participation was identified as important across four phases: planning, indicator design and use, data collation, and evaluation use. It was found that primary stakeholders may be involved in the indicator design and use phase; but are not routinely included in the planning, data collation and evaluation use phases. Findings from the eight case examples pointed to positive impacts on indicator design including the use of culturally appropriate and contextually relevant indicators, as well as participant empowerment and engagement in evaluation practice. Other findings highlighted that lack of skilled facilitators, the interplay of power dynamics and the length of time participatory evaluation practice takes may have negative impact on the engagement of primary stakeholders in the evaluative processes.

The practice of participatory indicator design is seemingly not widespread in the field of international development. While there is literature to be found that can provide some guidance on participatory evaluation practice, including design of outcome indicators, it seems that individual organisations reinvent processes for engagement on a case by case basis. The proposal is mooted that systematic capability building across the NGO sector that includes exploration on how the four phases of participatory evaluation practice can be built into organisational processes is required.

Acknowledgments

Oh my gosh! It's done. There are so many people who have supported me to get to the point of completion – thank you all.

Firstly I have to thank Dr. Gerard Prinsen and Associate Professor Robin Peace who have encouraged me and provided practical advice throughout. I could not have done this without your generous input. Gerard I loved reading the sarcastic and crazy comments you wrote in the margins – they made me smile and sometimes want to make you feel some pain. Robin you were so positive and made me feel like I could, even when I was down – thank you. Hopefully this thesis reflects the amount of time and energy you both threw into it. If I don't get a good grade, it's probably a reflection on … me. Staff of the Development Studies Department – Regina, Maria, Rochelle and Glenn – we all know this thesis would not be complete had it not been for the writing retreat and the discussions we had there – thank you.

My friends, hopefully you understood my withdrawal from all things social. Thank you for your regular enquiries into how the thesis was going. Thank you also for the constant jokes at my expense, like my postgrad diplomas not being real qualifications – top-ups as you called them. This one's for real! I should also extend my gratitude in advance to you all for the surprise graduation party that you're already organising with the kids – you're all so special!

Felix and Prisca, my kids, whose lives have been entertainment breaks during the writing of this thesis. This one's for you. Thank you for inspiring me to lead by example. Now that this is done I'll be hanging out with you and your friends constantly – I'm sure you'll love that. Naani, Nyasha and Hezvo, thank you for your belief in me.

Gerard, my friend. I apologise for not getting the word 'moist' into this thesis. Do a PhD and get it into yours. I did manage posit and reflexivity so hope that counts for something. You came into my life at a time when I was getting into my thesis – and I lost momentum – thanks honey. I love you and I'm thankful for the practical support you've given.

I would like to thank God for being there and giving me the strength and stamina needed to complete this thesis. I owe you Lord.

Lastly, the women, both staff and clients of Organisation X. Thank you for allowing me to do my research with you. To staff at Organisation X, a big thank you for being open with me and

for not withholding any information that was useful for this study. To the clients of Organisation X, thank you for the honest conversations we had. You allowed me into your lives and inspired me by your strength and courage to want more for yourselves. The words of one woman particularly continue to inspire me and enhance my interest and enthusiasm for participatory evaluation practice. Your words, said during an interview, and unedited are presented on the next page in the form of a poem.

Measure

*How do you measure anything
if you're not a part of it?
If you're not getting right in there amongst it
and
actually knowing as much as you possibly can
to evaluate anything?
You've got to be in it!
You've got to understand it!
You've got to look at it from every aspect!
Pull it apart as many times as you might need to
otherwise you fall into assumption,
and what you believe,
and not what's really best for the thing that you're pulling apart!*

*And I guess too
that when there's consumers involved
everyone has a different take
because
everyone leads a different life
and everyone has a different purpose.
So the importance of having the consumer involved
is huge
and I know that's often lacking
in terms of evaluation and research.
They'll bring that at the very end of it,
instead
of at the very beginning.*

*And I think it would change the whole result.
A lot of the consumers fill out the evaluation
and sign it
because that is what they are told to do.*

*When you begin any process,
and you're at the start,
and you're in the middle
and you come through the end,
you have a different value to it.
So it's different.
So knowing me I would be involved with everything.*

Local case study participant

Table of contents

List of figures and tables

Figures

Tables

Glossary

Hapu	Clan
Iwi	Tribe
Mana	Honour, authority
Mana whenua	Traditional authority exercised by iwi or hapu in an identified area.
NGO	Non-governmental organisation
Outcome indicators:	This thesis uses the term 'outcome indicators' to cover both outcomes and impacts. A working definition used for the purpose of this thesis, based on the premise that development interventions seek to bring about positive change, is: outcome indicators are qualitative and/or quantitative measures of intended positive change brought about as a result of an intervention.
Primary Stakeholders:	The people at whom a development intervention is targeted. For instance the primary stakeholders for literacy and numeracy classes will be men and women who attend. This does not take into account members of staff of the organisation who deliver these programmes.
RBA	Results-Based Accountability
SMART	Specific, Measurable, Achievable, Relevant, Time-bound
SPICED	Subjective, Participatory, Interpreted and communicable, Cross-checked and compared, Empowering, Diverse and disaggregated.
Taha hinekaro	Emotional wellbeing
Taha tinana	Physical wellbeing
Taha wairua	Spiritual wellbeing
Taha whanau	Social wellbeing

UNDP   United Nations Development Programme

Waiora   Wellbeing

Whanau   Family or families

Chapter One: Introduction

Organisations involved in implementing development interventions find there is pressure on them to show evidence of the efficacy of their work (Thayer & Fine, 2001). However, the difficulties associated with evaluating the value, merit and worth of such interventions have caused much discussion and debate about how and by whom outcomes should be determined (Holte-McKenzie, Forde, & Theobald, 2006; Parkinson, 2009; Vernooy, Qiu, & Jianchu, 2006). Although evaluation has, in the past been the domain of 'experts' coming mostly from developed countries, and using quantitative methods, interest in participatory processes has grown, as has the interest is measuring qualitative aspects of interventions (Estrella & Gaventa, 1998). From conventional monitoring and evaluations that were and still are expert-led, there has been growth in methodologies for conducting participatory evaluations that involve those who benefit from development interventions, including in the determination of outcome indicators (Vernooy et al., 2006).

The reasons for including primary stakeholders in the design of outcome indicators and evaluations are numerous. Some of these reasons are found in international case examples used as part of this research study. These include: involving primary stakeholders in evaluation processes in order to facilitate organisational learning and development to ensure that learning is embedded as part of the organisation's cycle (Parkinson, 2009); to build the capacity of local stakeholders to conduct participatory and self-monitoring and evaluation for the projects and programmes they are involved in through local organisations (Hamilton et al., 2000; Holte-McKenzie et al., 2006; McDuff, 2001; Vernooy et al., 2006); to empower primary stakeholders (Holte-McKenzie et al., 2006; Parkinson, 2009; Vernooy et al., 2006); to build the capacity of research teams in participatory monitoring and evaluation (Njuki, Mapila, Kaaria, & Magombo, 2008; Vernooy et al., 2006); to fulfil donor requirements (Izurieta et al., 2011; Scheyvens, 2007); and to increase utilisation of evaluation findings by the organisation and primary stakeholders (McDuff, 2001).

This first chapter introduces the research topic, stating its aim and questions, and how the research was conducted. The motivation and positionality of the researcher is outlined. The final section of the introduction outlines the roadmap of the wider thesis.

Research aim and questions

Situated largely within participatory evaluation research in international development, the aim of this research is "**to examine the involvement of primary stakeholders in the design and use of outcome indicators, and evaluation use for development interventions**." The research questions associated with this aim are:

Question 1: At what phases in participatory evaluation practice are primary stakeholders involved?

Question 2: What are the characteristics of outcome indicators designed and used by and with primary stakeholders?

Question 3: What are the dynamics of involving primary stakeholders in design and use of outcome indicators, and evaluation use, and the lessons learned? and

Question 4: What is achieved when primary stakeholders are involved in the design and use of outcome indicators, and evaluation use?

Theoretical and methodological frameworks are used in this research to meet the above aim and to answer the research questions. My research explores the theoretical framing of 'participatory' evaluation approaches that ostensibly involve primary stakeholders, an analysis of international case examples that offer practical application of the concepts, and empirically explores a local case study in which primary stakeholders were recruited to help develop evaluative indicators. Both the theoretical and empirical approaches investigate the phases in participatory evaluation practice at which primary stakeholders are involved, the characteristics of the indicators designed, the dynamics of primary stakeholder involvement, and what gets achieved.

The research is framed in the context of evaluation theory, that is, it draws on insights from four methods used to conduct this study: a descriptive review of international case examples, document analysis, participant observation and semi-structured interviews. Thematic analysis is used to find meaning in the data and generate discussion.

For the purpose of this research, a working definition for primary stakeholders is 'the people at whom a development intervention is targeted.' For instance the primary

stakeholders for literacy and numeracy classes will be men and women who attend. This does not take into account members of staff of the organisation who deliver these programmes.

This thesis uses the term 'outcome indicators' to cover both outcomes and impacts. A working definition used for the purpose of this thesis, based on the premise that development interventions seek to bring about positive change, is outcome indicators are qualitative and/or quantitative measures of intended positive change brought about as a result of an intervention.

This research is borne out of interest arising from my personal experiences.

Motivation and positionality

While one objective of this thesis is to partially fulfil the academic requirements of a masters qualification, there is also my own personal motivation and subjectivity.

My interest in participation of primary stakeholders, particularly in the design and use of outcome indicators, are intertwined with personal experiences. I am a Zimbabwean woman who grew up in an orphanage in Zimbabwe. Until the age of 19, when I left the orphanage, I was not able to state what I wanted in life and how I wanted to live that young life. My progress in and quality of life was measured by, amongst other things, my ability to communicate fluently in English, my school performance and the number of times I had to see my social worker, rather than by relationships with other children and those who I interacted with outside the large fence that defined the parameters of my existence. Later in life I would question whether the British and Irish matrons who raised us had thought about how their decisions, for instance not allowing us to speak our own languages, would affect us later in life.

My first job was as a secretary with a development organisation, the Zimbabwe Council of Churches that focused on development and issues of justice, peace and advocacy. My work involved attending meeting with funding partners to discuss the work we were doing and some of the funding challenges the organisation faced. Our funding partners often raised two issues with us. One was the need to evaluate the work we were doing; the other was to change the focus of our programmes, as the state of the nation evolved, even though we felt the programmes being run at the time met the changing context. My personal

frustrations were twofold. Firstly, evaluations were donor-driven and had to be done to the donors' timetables. Measures used were often those of our funding partners. I remember wondering if the focus on the number of bags of cement and other materials that went into construction of rainwater harvesting tanks was more important than what the communities told us of the benefits of immuno-compromised relatives having access to clean water close to home. As staff we convened a meeting and decided that because we had been hearing whispers from the donor world about logframes, we would lead the process ourselves before they imposed them on us. My second frustration was the difficulties we faced in trying to design indicators appropriate for measuring the outcomes of development projects we supported. I wonder if the indicators we designed and used might have been different had primary stakeholders been at the table to say what needed to be measured, because it was their lives we were keen to improve.

Since coming to New Zealand I have been involved in two evaluations overseas. The first was in my home country, Zimbabwe, and was commissioned by a European international non-governmental organisation (NGO). The evaluation, of a drought-relief and HIV/AIDS integrated project, was conducted by a Dutch evaluator and me. The Zimbabwean NGO worked hard to achieve the change the indicators tracked, indicators that were designed by the resident donor representative using the Logical Framework Approach. The evaluation questions were drawn up by the European international NGO and the Zimbabwean NGO was given the opportunity to comment and give their input. The second evaluation was in Rwanda and was commissioned by another European donor organisation, and conducted by the same Dutchman and myself. The evaluation questions were framed by the donor, and the Rwandan organisation being evaluated was given a chance to comment on these. I remember thinking at the time that had the Rwandan NGO commissioned the evaluation the questions used in the evaluation would have been different, and our findings may have been more useful for the organisation. As it was, my feeling was the evaluation findings were known to them and therefore had little use. My experiences in both cases, and my personal interest in indicators, led me to do postgraduate studies in Development Studies and in Social Sector Evaluation Research.

Working now within the development sector in New Zealand, I am interested in and work with local organisations to build capacity on monitoring and evaluation. While this is a very small part of my work, I try to focus on this because very few organisations I work with

have monitoring and evaluation plans and do not seem to evaluate their work. Evaluation is done only to meet the requirements of funding organisations using, for instance, Results-Based Accountability with focus on two to three indicators contained in the funding application.

Based on my experience in Zimbabwe and New Zealand, my interest in how the voice of the poor and voiceless can and should be heard has been heightened. The aim of examining the involvement of primary stakeholders in the design and use of outcome indicators and evaluation use for development interventions has thus been influenced by my personal journey of being a child and young adult without space to articulate my wishes, and later being in a position where I should have been more mindful of the voices of the voiceless in my work in Zimbabwe. Trying to incorporate the voices of primary stakeholders by use of appreciative inquiry to evaluate work in Zimbabwe and Rwanda, and my desire to influence involvement of primary stakeholders in evaluation processes, have also led to the aim of this thesis.

Because I come from an oral culture in which we make sense of life by being with one another and conveying messages through stories, I was fascinated by case study research in that I felt it would provide me with room to use different research methods to explore my topic. I opted for semi-structured interviews because I felt research participants would be more engaged in a one-on-one conversation. It was important to me that the start of the interview provide them with an opportunity to talk about what they wanted to, in relation to the process they were going through with the local case study used in one aspect of this research. Being mindful that my personal interests and motivation could get in the way of valid and credible research, the research uses a case study approach with four methods to be triangulated to reduce bias.

Thesis structure

This thesis is divided into seven chapters.

Chapter One has introduced the research topic and its aim and questions. It has briefly outlined the context within which this research is conducted and the methods used. The motivation and positionality of the researcher have been presented.

Chapter Two reviews the theoretical context and commitments of participatory evaluation through an overview of literature that offers definition and explanation of participatory practices, points to particular participatory approaches and rehearses some of the guidance offered in relation to indicator design. It includes a rubric developed as part of this study to analyse the internal case examples and the local case study.

Chapter Three focuses on methodological considerations taken into account for this study. It begins by stating the philosophy that underpinned the research, and discusses the methodological approach used in conducting the literature review, selecting the eight case examples analysed in Chapter Four, and an empirical project based on a case study of a local NGO. Fieldwork preparations and experiences, including ethical considerations and reflexivity are presented. The chapter draws to a close with discussion on data analysis and coding.

Chapter Four explores eight international case examples that were identified to demonstrate cases where primary stakeholders have been involved mainly in the design and use of outcome indicators for development interventions. It is divided into four sections to address the research questions.

Chapter Five is a presentation of the local case study. The context of the local case study is provided here, to avoid any confusion that the context applies to the international case examples. The chapter is divided into two parts. Part One sets out the New Zealand context within which case study research was conducted. Apart from outlining indicator design and use in New Zealand, it gives a brief overview of the organisation used as a case study. Part Two presents the fieldwork findings following, as closely as possible, the typology of the rubric formulated in Chapter Two, and the presentation of the eight case examples.

Chapter Six is discussion on the findings and conclusion. It discusses the main themes drawn from the international case examples reviewed in Chapter Four and the local case study in Chapter Five. It deliberates on the phases of involvement in participatory evaluation practice, characteristics of indicators designs, the dynamics of involving primary stakeholders, and what is achieved by the case examples, and in particular the local case study. The chapter concludes the thesis by putting forward the key themes that emerged in the attempt to achieve the aim of the research. It highlights issues that need to be further explored and the limitations of this study.

Chapter Two: Literature review

Introduction

This chapter provides background to the involvement of primary stakeholders in evaluative activities by reviewing literature on the theoretical concepts and commitments of participatory evaluation for development interventions. It sets the scene for Chapter Four that directly examines a set of international case examples, and a local case study in Chapter Five in which primary stakeholders have been involved in developing evaluative indicators.

In the first section, the broad field of participatory evaluation is considered, followed by a more detailed discussion of what outcome indicators are. Their use in the field of international development is presented in the next section that also provides discussion on some particular approaches within participatory evaluation that focus on indicator design. The following section reviews guidance provided in the literature in relation to indicator design within participatory frameworks. Such guidance provided scope for the design of a rubric, presented in the last section, to be used in retrospective classification of existing international case examples and for evaluation of the local case study.

Theoretical concepts

Traditionally, organisations involved in development interventions ranging from child nutrition to adult literacy have been funded by international NGOs, and multilateral and bilateral donors who commission evaluations for various reasons including public accountability. This has led to NGOs viewing evaluations as donor-related activities that are designed for accountability and attracting funding and that are implemented by experts contracted by development agencies. The research and evaluation methods used by external evaluators have been questioned, as they are often not deemed appropriate for the communities within which evaluations are carried out (Rist, 1995).

Marsden and Oakley (1990) suggest that methods that are participatory and based on the idea of self-evaluation should be used for evaluating development interventions. They assert that the principles and practice of these methods are in keeping with regular on-going activities of development interventions; are cost-effective and systematic; and take into

account the capacity and availability of community members. Evaluation approaches that accommodate such thinking include empowerment, participatory, utilization-focussed, stakeholder-involving and collaborative evaluations. These approaches, apart from challenging donor perceptions and practices of evaluations that are largely expert-led, seek to involve a range of stakeholders in evaluative thinking and practice.

Participatory evaluation – what is it?

Participatory evaluation methodology has its roots in community and international development. It arose as response to traditional evaluation practice to give voice to primary stakeholders' perspectives of development interventions, a domain that had been controlled and dominated by international donors (Brunner & Guzman, 1989). Cousins and Earl (1992, pp. 399-400) define participatory evaluation as:

> Applied social research that involves a partnership between trained evaluation personnel *and* practice-based decision-makers, organization members with program responsibility, or people with a vital interest in the program.

In the late 1970s participatory evaluation approaches emerged in importance in international development in response to the perceived incongruity between the needs of primary stakeholders and development interventions targeted at addressing these needs (Cullen, Coryn, & Rugh, 2011). Participatory evaluation became a mechanism through which the values, needs and aspirations of primary and other local stakeholders were recognised. Abma and Widdershoven (2008) contend that evaluations that involve primary stakeholders are achieved when conversations are held between the evaluator and other local stakeholders, and relationships stem from such dialogue and inclusive practice. Another driver for primary stakeholder inclusion has been the apparent unequal distribution of power that does not favour primary stakeholders (Vernooy et al., 2006).

Participatory evaluation places emphasis on the need for relationships to be built between the evaluator and local stakeholders (Cousins & Whitmore, 1998). The nature of the relationship needs to be such that various stakeholders feel they can trust the evaluation process and be involved meaningfully. The implication is that an evaluator, or researcher, works alongside those who are directly affected by development interventions, that is, the programme staff and management, community members and leaders, individuals, and groups within a particular community. Cousins and Earl (1992) state that local stakeholders should

be involved in all aspects of the evaluation process from defining the parameters of the evaluation to disseminating the findings.

The benefits derived from conducting evaluation in such a relational manner include access to knowledge of the local context (Chambers, 1997a), awareness of cultural considerations (Tuhiwai-Smith, 1999), and knowing what data to collect in order to measure what's important (Wallace, 2010). Cousins and Chouinard (2012) speak of the complementarity of the evaluator's professional knowledge and practice, and the depth of knowledge and understanding of context and programmes that go into participatory processes. Pretty and Vodouhê (1998) write about participatory learning, in which learning occurs as different stakeholder groups work alongside each other. Such process allows for different stakeholders to be cognisant of the varied, and at times divergent, meanings placed on concepts such as sustainability, participation, and development. Another advantage of participatory evaluation is that, when involved in the planning and design of indicators, primary stakeholders are also in charge of the process and their views inform what is to be done (Kibuga, Bibby, & Alfred, 1999; Vernooy et al., 2006). Such inclusive processes also increase the use of evaluation findings because there is local ownership and understanding of what is being proposed (Patton, 2008).

Some evaluators further differentiate participatory evaluative approaches such as that made between practical participatory evaluation and transformative participatory evaluation. Cousins and Whitmore (1998) state that the goal of practical participatory evaluation is underpinned by a philosophy that believes in the social nature of knowledge and its construction. It rejects the top-down approach typical of research and development that involves a limited number of stakeholders in some aspects of evaluation processes in order to increase utilisation and contribute to decision-making. Practical participatory evaluation "functions to support program or organisational decision making and problem solving" (Cousins & Chouinard, 2012, pp. 22-23). Transformative participatory evaluation is concerned with giving voice to the marginalised, usually primary stakeholders or those affected by development and other interventions, in evaluation practice. It focuses on changing social structures by empowering grassroots communities and challenges the notion of objectivity "by making explicit the political connections of knowledge, power and control" (Brisolara, 1998, p. 29). Transformative participatory evaluation is similar to Participatory Action Research and Participatory Rural Appraisal which address social inequalities (Weaver &

Cousins, 2001) and puts those whose lives are most affected by the evaluation at the centre of the process from beginning to end (Brisolara, 1998).

Culture is understood to have significant influence on research and evaluation processes. Trickett, Watts and Birman (as cited in Donaldson & Scriven, 2003, p. 168) posit that cultural consideration is "the improved understanding of diverse cultural groups . . . which can . . . improve our evaluation plans in the direction of greater social justice." During the various phases of an evaluation the epistemologies of evaluators, program staff and communities often differ as a result of their varying cultures, lived experience and ways of knowing. Lather (1991, p. 105) contends that "all researchers construct their object of inquiry out of the materials their culture provides." Because evaluations are not culture-free (Chouinard & Cousins, 2009), there are instances when it is appropriate to employ evaluators who have specific cultural knowledge of a community, especially when the culture is different from that of the evaluator (Lunt, Davidson, & McKegg, 2003; Tuhiwai-Smith, 1999).All participants bring their history, a way of knowing the world and their values, to the process (Kvale, 2006). From a constructivist perspective, reality is socially constructed and based on the way individuals view the world, their ways of knowing, and their knowledge systems.

Participatory evaluation, like all other approaches to evaluation, is not without its critics. While achieving participation is an important goal of participatory evaluation, it is not always desirable for all stakeholders involved. Cooke and Kothari (2001) warn about participation possibly being a tyranny and that methods used can get in the way of using others that may be more productive in attaining the aspirations sought through participation. Baur, Abma, and Widdershoven (2010) also caution against putting different stakeholder groups together without addressing power differentials, as this can lead to primary and marginalised stakeholders becoming even more silent and withdrawn from such processes.

Involving stakeholders in evaluations can also be costly and time consuming. Critics point out that it is difficult to plan the amount of time and resources required to conduct participatory evaluation in advance. Sufficient time has to be allocated for the various aspects of evaluation including learning, sharing, building capacity of local stakeholders and understanding the diverse perspectives (Crishna, 2006).

Participatory evaluation, however, is seen as important in the field of international development. Within the practice of participatory evaluation and research, it is acknowledged

that indicators continue to be vital for monitoring and evaluating programme success (Carruthers & Tinning, 2003; Guijt, 2000; Scheyvens, 2007; Vernooy et al., 2006). What is questioned is who decides on indicators, evaluation questions, and evaluation use.

Outcome indicators

The historical background of indicators as outlined by Prinsen (2013) traces their possible foundations to Drucker, who in the 1950s introduced the idea of 'management-by-objectives.' A consultant in the private sector, Drucker mooted the need for yardsticks to measure productivity (Drucker, 1955). This thinking spread to public administration when, in the early 1980s, the use of performance indicators was introduced in Britain and spread from a few indicators to over 1800 in about five years (Doreen, Buckaert, & Halligan, 2010). The indicator-led practices became central to New Public Management.

As the importance and practice of New Public Management was spreading in public administration, multilateral and bilateral organisations, such as the United Nations, World Bank, and the UK Department for International Development also adopted the use of indicators. In 1990, for instance, the United Nations designed hundreds of indicators for its Human Development Index (UNDP, 1990), the first in a series of many to follow. The practice of indicator design in the field of development spread rapidly. Banks of indicators from a variety of fields within development include: World Bank indicators for development areas including agriculture and rural development, education, health, trade, gender and environment (World Bank, n.d.); the United Nation's Millennium Development Goals whose eight goals have 48 indicators (United Nations, n.d.); Community Indicators Consortium that has about 300 listed community indicator projects (Community Indicators Consortium, n.d.); and as a New Zealand example, Community Waitakere that lists indicators for nine domains in community and social wellbeing (Community Waitakere, 2012b).

Definitions of outcome indicators

There is no one agreed definition of outcome indicators. (Mathison, 2005, p. 288) defines outcomes as "changes, results and impacts that may be short or long term; proximal or distal; primary or secondary; intended or unintended; positive or negative; and singular, multiple or hierarchical." The United Nations Development Programme (UNDP) (2002) describes an indicator as a "signal that reveals progress or lack thereof towards objectives;

means of measuring what actually happens against what has been planned in terms of quantity, quality and timeliness."

UNDP (2009, p. 67) describes outcome indicators as "progress against specified outcomes. They help verify that the intended positive change in the development situation has actually taken place." Save the Children UK (2008, p. 7) states that outcome indicators measure the "intermediate changes as a result of the project or programme activities … tells us if activities are bringing about the intended … changes." According to Statistics New Zealand (n.d, p. 8) an indicator is a "summary measure related to a key issue or phenomenon that can be used to show positive or negative change. The evaluative nature of an indicator distinguishes it from the descriptive nature of statistics". Unlike indicators for inputs, activities and outputs, outcome and impact indicators are measures for high-level goals. At this level, outcome and impact indicators assist in the identification of the contribution that an intervention has made as ascertaining clear causal relationships can be challenging (DeGroff, Schooley, Chapel, & Poister, 2010). Because of the lack of consistency in terminology there is often a tendency for outcome and impact level indicators to be confused or used interchangeably.

Direct and proxy

It is important to note that although there are aspects of an intervention that indicators can measure directly, there are some aspects that are more difficult. Indicators related to outcomes and impacts of interventions are based on the consequences of the intervention rather than the intervention itself and are thus said to be proxy (UNDP, n.d). To clarify this insight, the United Nations Development Programme uses the poem by Christina Rosetti:

> Who has seen the wind?
> Neither you nor I.
> But where the trees bow down their heads,
> The wind is passing by! (UNDP, n.d, p. 2)

Although the wind cannot be seen, an indicator that it is blowing is the bowing treetops. A word of caution is given in that while indicators are useful in providing information about what has occurred, they cannot tell us why that change has taken place or why the intervention has made a difference.

Planning, process and product indicators

Donabedian (1998) asserts that indicators can be developed under three categories namely structure, process and outcome, where structure is concerned primarily with the inputs for a programme (for example personnel, money, facilities), process is to do with what happens during delivery of an intervention, and outcomes comprise the result of the intervention. This is in keeping with what Jacobson (1991) proposes when he outlines the need for indicators at planning, process and product level.

Quality of indicators

The quality of indicators is ascertained by their ability and appropriateness to produce reliable and valid data in relation to a desired outcome. Scheyvens (2007) describes some of the challenges of designing appropriate indicators for monitoring and evaluation of development programmes including that, sometimes, the indicators become a target at risk of overlooking more important unintended outcomes, they may not be culturally appropriate, or that qualitative indicators, often seen as inferior to quantitative ones, can provide more accurate measures.

Marr (2009) proposes two templates he considers useful in the design of indicators. In the first instance he proposes a 10-step performance indicator decision framework that, through answering questions, leads to the design of high performing indicators. The second template assists in the design of such indicators by going through a 20-step process (see Appendix One).

When judging the quality of the indicators designed for development interventions in a participatory manner Roche (1999) suggests that SMART (Table 2.1) and SPICED (Table 2.2) are two different criteria that can be used.

Table 2.1: SMART properties of indicators

SMART properties of indicators	
Properties	**Definition**
Specific	Indicators should reflect those things the project intends to change, avoiding measures that are largely subject to external influences
Measurable and unambiguous	Indicators must be precisely defined so that their measurement and interpretation is unambiguous Indicators should give objective data, independent of who is collecting the data Indicators should be comparable across groups, projects thus allowing changes to be compared and aggregated
Attainable and sensitive	Indicators should be achievable by the project and therefore sensitive to changes the project wishes to make
Relevant and easy to collect	It must be feasible to collect data on the chosen indicators within a reasonable time and at a reasonable cost Indicators should be relevant to the project in question
Time-bound	Indicators should describe by when a certain change is expected

Source: Roche, 1999 p48.

Table 2.2: SPICED properties of indicator development and assessment

SPICED properties of indicator development and assessment	
Properties	**Definition**
Subjective	Informants have a special position or experience that gives them unique insights which may yield a very high return on the investigators times. In this sense, what may be seen by others as 'anecdotal' becomes critical data because of the source's value.
Participatory	Indicators should be developed together with those best placed to assess them. This means involving a project's ultimate beneficiaries, but it can also mean involving local staff and other stakeholders.
Interpreted and communicable	Locally defined indicators may not mean much to other stakeholders, so they often need to be explained.
Cross-checked and compared	The validity of assessment needs to be cross-checked, by comparing different indicators and progress, and by using different informants, methods and researchers.
Empowering	The process of setting and assessing indicators should be empowering in itself and allow groups and individuals to reflect critically on their changing situation.
Diverse and disaggregated	There should be a deliberate effort to seek out different indicators from a range of groups, especially men and women. This information needs to be recorded in such a way that these differences can be assessed over time.

Source: Roche, 1999 p49.

When the objective of designing indicators is to verify and plan for an intervention, then it is appropriate to use the SMART criteria. The SPICED criteria is more applicable when designing qualitative proxy indicators in participatory processes (Roche, 1999, p. 49). Roche (1999) suggests that the two criteria can be combined.

The Scottish Government (n.d., Measures and indicators Section, para. 10) suggests that indicators are useful when they give information that helps to establish "How much has changed? [quantity]; How beneficial has the change been? [quality]; Who has benefitted and who has not? [equity]; What resources have been used? [efficiency]; and How far have the planned outcomes been achieved? [effectiveness]."

Indicator design approaches used for development interventions

Within the practice of participatory evaluation and participatory research, two questions have been raised by Chambers (1997b), and Estrella and Gaventa (1998) in their books aptly entitled *Whose Reality Counts* and *Who Counts Reality,* respectively. The first question, pertaining to whose reality counts, points to the politically charged nature of setting evaluation questions and indicators for development interventions, and can be answered by examining who has been involved in the design of evaluations and indicators for an intervention. It is rare for multiple stakeholders of a project to design the same indicators for a development intervention as they may have different information needs. The second question, who counts reality, can be known by finding out who designs evaluation questions and makes recommendations for change.

Outcome indicators have traditionally been designed by experts in a top-down manner for use with the Logical Framework Approach in what has become known as Results-Based Management (Fraser, Dougill, Mabee, Reed, & McAlpine, 2006). As pushback to this expert-led way of designing indicators, participatory approaches such as Participatory Rural Appraisal and participatory monitoring and evaluation became ways of involving primary stakeholders. With growth in participatory evaluation theories and practice, Most Significant Change (Dart & Davies, 2003), as an indicator-less evaluation method, was introduced to offer another alternative to top-down approaches of evaluation.

Logical Framework Approach

The Logical Framework Approach is the most used approach to planning and management of development interventions. An example of its continued use is the Department for International Development's instruction that states:

> From January 2011 Departments must use the new DFID Business Case <u>for all newly approved projects.</u> Since the logframe is an integral element of the Business Case all newly approved projects regardless of project value must also now contain a logframe (2011, p. 1).

The logframe, the matrix-based visualisation of the Logical Framework Approach, is almost always a requirement by donor agencies, especially for large-scale projects seeking funding and other support. The approach uses logic to link activities, outputs and purpose, their associated indicators, means of verification and assumptions to a single goal. Complex

realities are simplified and put into logical sequences in which indicators can be used to trace an intervention's success (Jacobs, Barnett, & Ponsford, 2010). Logical Framework Approach has its roots in corporate and military planning contexts from the 1960s, and is straightforward to use in these contexts, where there is usually one objective such as gaining large profit margins or national security. This is not the case in many development interventions where there are multiple goals, borne from competing interest of numerous stakeholders, some of who are negatively affected by the intervention (Gasper, 2000).

Proponents of the Logical Framework Approach state that all stakeholders can and should be involved in all aspects of this approach including designing indicators (Bakewell & Garbutt, 2005; Department for International Development, 2011). Yet the time and effort required to develop the initial logframe, using participatory processes, that entail negotiating views of various stakeholders for consensus building, is such that the idea of revising the matrix on an ongoing basis in a participatory manner as the intervention progresses is difficult. This sentiment was expressed by an international development NGO during a stocktake of the Logical Framework Approach and its use:

> The process of participatory logframe planning is too idealised for my liking. It may take us years to help diverse communities reach consensus about priority, higher risk issues. Only then would a participatory logframe planning approach become relevant. To pretend that rapid participatory planning can somehow replace this risks being superficial and promoting the inclusion of local elites rather than poor people (Bakewell & Garbutt, 2005, p. 7).

Some argue that it uses a top-down approach (Gasper, 2000; Jacobs et al., 2010; Levermore, 2011), and is usually created post-project design, thereby failing to guide logical thinking through the design phase (Gasper, 2000). Gasper (2000, p. 19) argues that *"the value of LFA* [Logical Framework Approach] *declines as we move from project design through to post-implementation evaluation"*. Because the logframe is a requirement and is then used to measure the success of interventions, the indicators can become a target for the organisation rather than a possible indication of success (Hummelbrunner, 2010). Logframes assume that those who devise indicators have perfect foresight of what is to come (Springer-Heinze, Hartwich, Henderson, Horton, & Minde, 2003). Pushback regarding the use of Logical Framework Approach came in the form of participatory monitoring and evaluation, including Participatory Rural Appraisal.

Participatory Rural Appraisal

Participatory Rural Appraisal has its roots in the agricultural sector. Emerging from rapid rural appraisal which seeks to mobilise local knowledge for analysis by experts somewhere else, Participatory Rural Appraisal seeks to give control to local people to manage their own development interventions (Chambers, 1994). Chambers (1994) posits that participatory rural appraisal, a collection of approaches and methods that have grown and spread since the 1990s, "Enable[s] local (rural or urban) people to express, enhance, share and analyse their knowledge of life and conditions to plan and to act" (p. 1253).

It is an approach to planning, monitoring and evaluation that focuses on involving local people in all aspects or research and evaluative processes, including in the design of indicators. It is based on participatory action and uses several methods including transect walks, pairwise ranking, wealth ranking, seasonal profiles, and asset mapping (Pretty & Vodouhê, 1998). Chambers (1997b) asserts that the strength of Participatory Rural Appraisal is that it is grounded in the voices of local people and informed by their wisdom. He suggests that Participatory Rural Appraisal processes result in building the capacity of communities to monitor and evaluate the work they do in a manner that leads to self-determined and community-driven action (Chambers, 1997b).

Criticism levelled against Participatory Rural Appraisal is that, although in principle the approach seeks to hear the voices and draw on the wisdom of the poor and vulnerable in communities, there is no guarantee that it is these people who will actively participate in the process (Aune, 2000). Without empowerment the poor will not be able to articulate what they would like to communicate and the status quo remains. Unlike the Logical Framework Approach, Participatory Rural Appraisal does not have a clear structure making it difficult to understand what the project is about (Aune, 2000). Critics also point out that not all people have skills in research and evaluation, and mobilising the poor does not necessarily cause these skills to be present in all. As Cleaver (2001, p. 46) holds that "development practitioners excel in perpetuating the myth that communities are capable of anything, that all that is required is sufficient mobilization … and the latent capacities of the community will be unleashed in the interest of development".

Other participatory approaches used in international development fall under the umbrella of participatory monitoring and evaluation.

Participatory monitoring and evaluation

Development of participatory monitoring and evaluation methods came about as a result of frustration with Logical Framework Approach and other approaches, that seemingly did not take into account local knowledge and grassroots development aspirations (Jacobs et al., 2010). Participatory monitoring and evaluation is a variant of participatory evaluation that explicitly includes monitoring, and assists in addressing power differentials between varying stakeholders. The assumption is that working together on monitoring and evaluation plans leads to shared goals in terms of what the evaluation process will achieve (Izurieta et al., 2011; Parkinson, 2009).

In keeping with increasing participatory practice in development, participatory monitoring and evaluation involves grassroots communities in decisions about how intervention success will be measured, and design of indicators that are contextually relevant (Estrella & Gaventa, 1998). Campilan (2000) argues that involvement of a variety of stakeholders in the process of monitoring and evaluation does not make it participatory. Instead, the participation in decision making and planning pertaining to which stakeholders should participate and in what way, and in varying roles results in a participatory monitoring and evaluation approach. Guijt (2000) proposes that participatory monitoring and evaluation involves 10 steps, one of which is the identification of indicators that provide the required information. The suggestion offered by Guijt (2000) is that stakeholders who should be involved, and want to be part of the evaluation process, should be identified, and that such stakeholders should decide how they would like to be involved throughout the process.

Participatory indicator design is one of the most challenging facets of participatory monitoring and evaluation processes (Guijt, 2000). This is largely due to the competing expectations, interests and needs that a diverse range of stakeholder groups and individuals bring to the negotiating process. Because each development intervention is carried out in a particular setting to address a particular need for a particular people, Pope and Mays (as cited in Holte-McKenzie et al., 2006, p. 51) assert that outcomes should seek to "represent that reality rather than to attain the truth."

Criticism has, however, been levelled at this approach as it is time consuming and can be very expensive to implement (Jacobs et al., 2010). While the aim is to include the realities of grassroots communities, it is usually the voices of the elite amongst them that are heard

and taken into account. There may also be tensions amongst the views and aspirations of grassroots communities and other stakeholders such as programme staff (Parkinson, 2009).

The usefulness of participatory monitoring and evaluation in measuring impact is sometimes questioned as the products of such processes, including indicators used, are not usually scientifically rigorous. To this, proponents of participatory evaluation point out that impact is culturally situated with those who have been affected by the development intervention (Chambers, 1997b). The premise of participatory evaluation is therefore that local stakeholders who have experienced change due to an intervention, and who are aware of what other factors may have influenced change, should participate in determining value, merit and worth of that intervention.

Vernooy et al. (2006) suggest that although indicators may be useful for measuring progress made during a development intervention, open-ended questions may result in more useful information. The Most Significant Change Technique offers such opportunity.

Most Significant Change, Indicator-less evaluation

Most Significant Change proposes a process through which stakeholders tell their stories and consider what has been the most important change in their lives as a result of a particular development intervention (Crawford & Bryce, 2003). Developed by Rick Davies and Jess Dart, Most Significant Change seeks to involve stakeholders in determining outcomes of an intervention from local perspectives (Dart & Davies, 2003). Like PRA, a benefit of using Most Significant Change, is its ability to engage more closely with grassroots communities and to solicit their views on the intervention. What can be problematic is that this is done in hindsight without baseline data (Willets & Crawford, 2007). Its reliance on storytelling after the fact means there need to be checks in place to verify stories and ensure accuracy, potentially making the process time-consuming.

The evaluation approaches used within the development context, albeit with their pros and cons, have a place in international development, and outcome indicators hold a place of importance for practitioners. With the growing popularity of participatory practice, the design of indicators is also a point of discussion in development.

Participatory design of indicators

Community-devised indicators are often contextually relevant and reflect the lived reality of particular communities (Fraser et al., 2006). Based on the discussions in available literature, it would appear that concerted effort has been put into participatory design of indicators for sustainability, resilience and sovereignty. Numerous stakeholders have been involved in a diverse range of processes with varying results. Such practice includes participatory design of sustainability indicators in South Africa (Terry, 2008)); forest management in Coastal British Columbia in Canada, (Fraser et al., 2006); understanding desertification in Botswana (Reed & Dougill, 2002); and assessment of community resilience to climate change (Elasha, Elhassan, Ahmed, & Zakieldin, 2005). The volumes of indicators that have been developed, especially within the field of sustainable development have led to the design of indicators being dubbed "an industry on its own" (C. King, Gunton, Freebairn, Coutts, & Webb, 2000, p. 631) Large amounts of time and money are utilised to develop reports on indicators, reports that are not used (Innes & Booher, 1999; C. King et al., 2000). C. King et al. (2000) propose that the indicator industry booms because of the reductionist manner in which the indicators are designed, resulting in their ineffectiveness or lack of use. Carruthers and Tinning (2003) corroborate this view in their observation that the farmers who participated in their research in Australia hardly recognised, let alone used, widely available sustainability indicators. They put forward that the missing element in the design of indicators are the end users.

General advice on indicator design is available, particularly within the field of sustainable development. The literature review for this thesis found little practical guidance on participatory design of indicators for development interventions.

Four phases of involvement in participatory evaluation practice

The literature reviewed points to a set of potential themes for organising data into phases that can be used when designing outcome indicators in participatory evaluation practices. Literature that was key to developing the themes include a literature review on participatory monitoring and evaluation conducted by Estrella and Gaventa (1998); a learning process for designing and using sustainability indicators (Reed, Fraser, & Dougill, 2006); advice on methodological considerations for participatory monitoring and evaluation (Guijt, 2000); steps in setting objectives and indicators (Lennie, Tacchi, Koirala, Wilmore, & Skuse, 2011); and a manual on participatory program evaluation (Aubel, 1999). The four themes

derived are planning, indicator design and use, data handling, and evaluation use. Each is considered in more detail below.

Planning

This phase focuses on identifying the diverse stakeholders who should be involved in the process, their needs, and what information is required. Campilan (2000), Freebairn and King (2003), and Guijt (2000) state that the identification of key stakeholders in the process should be a starting point. Aubel (1999) suggests setting up an evaluation team, made up of people from different stakeholder groups to guide the process. The different stakeholders will have varying degrees of knowledge and experience of participatory evaluation. N. Johnson, Lilja, Ashby, and Garcia (2004) advise that access and ability are cornerstones for building participatory monitoring and evaluation capacity. By access N. Johnson et al. (2004, pp. 217-218) mean "the opportunity to participate in a monitoring and evaluation process that includes more than one stakeholder group", with ability being "the skill or knowledge required to do something".

Because different stakeholders and stakeholder groups have diverse expectations, needs and interests, there is need to clarify what the information requirements are and how these will be realised (Guijt, 2000). Freebairn and King (2003) advise that different stakeholders should put forward what it is they would like to contribute to the process in line with their information needs. Carruthers and Tinning (2003) state that it is important for there to be clarity on who indicators are being designed for and what their use will be. They suggest that "the foremost question really is 'Who is asking for information, and what are they going to do with it?' This line of questioning forms the starting point for designing appropriate indicators" (p. 320). Decisions should also be made on the approaches and methods to be used for evaluations (Guijt, 2000; Reed et al., 2006). Once decisions have been made on the approach to be used, the process of designing indicators ensues.

Indicator design and use

This phase is about the design of indicators for a particular development intervention, and or its evaluation. In participatory evaluation, discussions about the nature of indicator design are usually held when the diversity of stakeholder groups come together to share and learn. Power dynamics are usually at play in such situations and Baur et al. (2010) suggest that the silent or marginalised in such processes are given the chance to communicate their

views before others. Ritchie (2000) proposes that if indicators are to be useful for primary and local stakeholders they need to reflect primary stakeholder experience of the phenomena for which indicators are being designed. (Antrop, 2000) concurs, suggesting that indicators designed to measure large scale change over time provide inadequate detail for local and primary stakeholders to use in decision making. It is, therefore, necessary for stakeholders to negotiate the indicators to be designed and selected if they are to be useful for all stakeholders (Guijt, 2000; Hamilton et al., 2000).

Indicators are for either strategic or operational use. Indicators for strategic use are often broad, to reflect their use at policy level, while indicators for operational use are narrow and specific, for decision-making on the ground or at programme level (Stoneham, Eigenraam, Ridley, & Barr, 2003). Both types of indicators are important and need to be designed based on their purpose and the context within which they will be used. If indicators do not lead to local benefits, Freebairn and King (2003) warn that the likelihood is that communities will not devote time or energy in collecting data.

Use of the indicators during pilot or actual evaluations should be done in a participatory manner (Estrella & Gaventa, 1998; Kibuga et al., 1999). In participatory evaluation practice primary stakeholders become more than data sources. They take on some roles evaluators usually perform. Alongside other stakeholder, they consider data collection methods to be used to ensure that they are commensurate with the data to be collected (Aubel, 1999; Guijt, 2000; Kibuga et al., 1999). There may be need to make changes to indicators or data collection methods after testing these or after conducting the first data collection task (Guijt, 2000; Reed et al., 2006).

Data handling

The data handling phase involves data analysis, reporting and dissemination of findings. As Guijt (2000) explains, once data has been collected, it needs to be analysed, organised into a report and disseminated to relevant stakeholders. Involving primary and other stakeholders in the process means that different epistemological interpretations of data are presented. Those who participate in the design and use of outcome indicators, will, as pointed out earlier, use their cultural knowledge to make sense of the object of enquiry (Lather, 1991).

There is, however, caution that if quantitative scientific rigour is sought in the process of data analysis or outcome indicator design, there may be tension between scientific rigour and participation. Reed et al. (2006) discuss this tension stating that there is often overlap between expert and community designed indicators, and that the balance can be found to accommodate both. Having evaluators or facilitators that know something about the topic for which indicators are being designed can also be useful for guiding communities in the design of indicators that are reliable (Hamilton et al., 2000).

Once the data has been analysed, the challenge becomes how the information is presented, as different stakeholder groups have different styles of communicating, levels of literacy and information needs (Guijt, 2000). Guijt (2000) states that this is possibly a core step in participatory monitoring and evaluation that is not well implemented. Hamilton et al. (2000) found that visual forms of documentation to record qualitative data increase use of the data collected.

Evaluation use

This phase is about ensuring that evaluation findings are useful and used by the organisation implementing the development intervention. One of the arguments put forward for involving primary stakeholders in the evaluation process is that this increases evaluation use (J. A. King & Ehlert, 2008; Patton, 2008; Somers, 2005; Thayer & Fine, 2001). Somers (2005) reports that extensive involvement of a variety of stakeholders in evaluation processes can result in changes to individuals, organisations and the intervention itself. He argues that the benefits go beyond the evaluation findings, as other outcomes can result from subsequent change. Thayer and Fine (2001, p. 106) found, through a survey of non-profit organisations, that evaluation findings are more likely to be used when the successes of interventions are documented, and when recommendations are made as it is useful to include "constructive criticism in order to modify programs". J. A. King and Ehlert (2008) posit that a plan of how stakeholders will continue to be involved in ongoing tasks is vital as this is likely to result in greater participation and use of the evaluation findings. Such inclusion of stakeholders in ongoing activities and what may be deemed by some as planning rather than evaluation use, become "a developmental process akin to action research" (J. A. King & Ehlert, 2008).

These four phases, planning, indicator design and use, data handling and evaluation use, provide a framework against which to assess or classify stakeholder involvement in participatory evaluation practice.

Classifying involvement of stakeholders in participatory evaluation practice

The guidance given for participatory design of indicators and the four themes that emerged above were used to devise a rubric against which to retrospectively classify some existing cases, where primary stakeholders have been involved in the design of indicators. This rubric is presented in Figure 2.1.

Figure 2.1: Rubric for assessing stakeholder involvement in participatory evaluation practice

Phase	Role	Extent Measure		
		Not really 0-3	Partially 4-6	Fully 7-10
1 – Planning	Initiators			
	Evaluation team members			
	Evaluation designers			
	Needs assessors			
	Trainers			
2 - Indicator Design & Use	Data collectors			
	Indicator designers			
	Indicators selectors			
	Indicator aggregators			
	Evaluators			
3 - Data Handling	Data analysts			
	Report writers			
	Information distributors			
4 – Evaluation Use	Evaluation users			

Source: Author.

In Chapters Four and Five, this rubric is used as a starting point to identify at what phases primary stakeholders have been involved in participatory evaluation practice in eight international case examples and one local case study.

Chapter summary

This chapter has presented the theoretical underpinnings of participatory evaluation and the approaches used for design of outcome indicators for development interventions within the practice of participatory evaluation.

The value that participatory methods add to evaluative processes of development interventions have been presented and include: presence of local knowledge in the design of indicators and methods to be used (Chambers, 1997b); joint learning (Pretty, 1995); bringing different epistemologies and interpretations of data into the process of analysis (Lather, 1991); and increased use of evaluation findings (Patton, 2008). While acknowledging the

significance and value of these benefits, the importance of ensuring that participation does not become a tyranny by giving primary and other stakeholders the option to not feel obliged to be involved (Cooke & Kothari, 2001) is also raised.

In the literature reviewed thus far, various writers give guidance on the process for conducting participatory evaluation, and participatory indicator design specifically. Of particular note is the methodological advice on practical aspects of participatory monitoring and evaluation based on a collection of cased studies that Guijt (2000) comments on. What is not found in the literature reviewed is a tool by which to assess participation of stakeholders in the design of outcome indicators and necessarily the processes that come before and after this task in participatory evaluation practice. Advice drawn from the literature, however, provides sufficient commonalities to determine four themes of relevance to understanding the scope of engagement of stakeholders across the full evaluation process. The four themes are: planning, indicator design and use, data handling, and evaluation use. To assist with addressing the research aim, these themes have been consolidated to create a rubric that is used for classification of stakeholder involvement in participatory indicator design, particularly in Chapters Four and Five.

Chapter Three: Methodology

Introduction

The aim of this research is to examine the involvement of primary stakeholders in the design and use of outcome indicators, and evaluation use for development interventions. A constructivist approach to research is undertaken and both case examples from international literature and an empirical local case study are used to answer the research questions.

This chapter presents the methodological considerations of this study. The first section presents the research questions and outlines the theoretical framework and rationale for a qualitative approach. The research is undertaken in three phases. The first phase entails a review of international literature on participatory evaluation theory and practice, and primary stakeholder involvement in outcome indicator design specifically. This material is set out in Chapter Two. The second phase also has an international focus but is designed to identify case examples where primary stakeholders were involved in indicator design for development interventions. These case examples and their analysis comprise Chapter Four. Phase three is the local empirical case study work undertaken in New Zealand that is described in Chapter Five. It explains the selection of the local case study and considerations of power relations and ethical framing. Fieldwork experiences, reflexivity, and the use of a research journal are discussed here. The last section examines the linkages between the two data collection phases and outlines how the data were analysed. The analysis of the second and third phase comprises chapter Six. Figure 3.1 provides an overview of the research process.

Figure 3.1: Research process

Literature Review
↓
Formulate research questions
↙ ↘

International case examples   *Local case study*
                                                Participant observation
Review secondary data   Semi-structured interviews
                                                Review secondary data

↘   ↙
Data coding and analysis
↓
Discussion
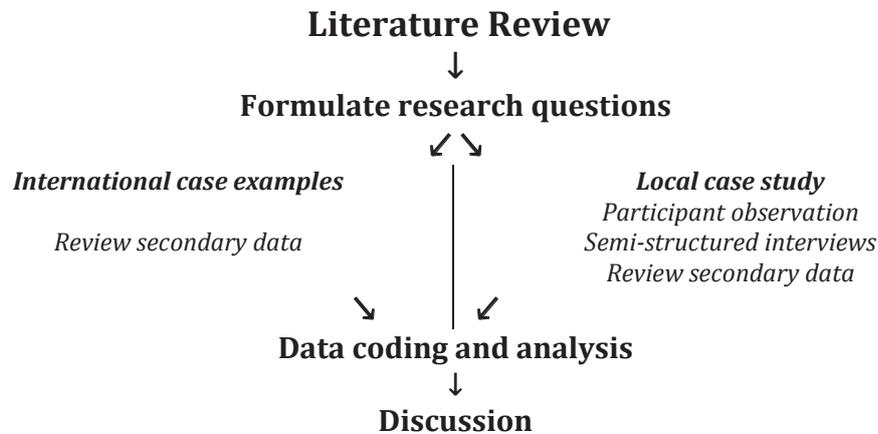
Source: Author.

Research questions and philosophical standpoint

The conceptual debates and discussion of involvement of primary stakeholders in the design and use of outcome indicators, and evaluation use, have been discussed in Chapter Two. In keeping with the research aim and questions, and to further explore themes discussed in the literature review, the research questions to be answered are:

Question 1: At what phases in participatory evaluation practice are primary stakeholders involved?

Question 2: What are the characteristics of outcome indicators designed and used by and with primary stakeholders?

Question 3: What are the dynamics of involving primary stakeholders in design and use of outcome indicators, and evaluation use, and the lessons learned? and

Question 4: What is achieved when primary stakeholders are involved in the design and use of outcome indicators, and evaluation use?

Question one is concerned with finding out the phases that primary stakeholders have been involved in evaluation practice, particularly the design of outcome indicators. Based on the themes extrapolated from the literature review and analysis of case examples the rubric, created and introduced in Chapter Two, is used to classify the phases of primary stakeholders' involvement in the case examples, and to some extent the local case study.

The second question seeks to gain insights into the nature of the outcome indicators designed by, and in some cases with, primary stakeholders. An examination of the case examples and the local case study is useful in devising a Venn diagram that can be used to depict the relationship between aspects of indicator design processes and the nature of the indicators.

The effects and challenges of working alongside primary stakeholders is explored in question three. This aspect of the research is largely informed by analysis of the practicalities of stakeholder involvement in evaluation processes.

The last research question brings to the fore some of the benefits of participatory indicator design processes, as reported in the literature review and found in analysis of the case examples and the local case study.

This inquiry was conducted within a constructivist paradigm. Use of a constructivist approach in social research is based on the premise that reality is socially constructed (Flick, 2004). It was important for me, as a participant researcher to be able to take part in and observe the different aspects of the primary stakeholders' processes in order to gain a deep understanding of their experiences (Brewer, 2000). It was also vital that the research be conducted in a manner that allowed participants to reflect on their experiences, and, through personal interpretation, place meaning on these. As such, my role as a researcher was to understand the differing interpretations and meanings placed on the numerous aspects of designing and using outcome indicators, and evaluation use by primary stakeholders.

Constructivism emphasises subjective interrelationships that exist between the researcher and the researched, stating that meaning is co-constructed (Hayes & Oppenheim, 1997). I was under no illusion that there would be multiple realities constructed by the individuals participating and that my own interpretations of reality would be intertwined in these. This is aligned with the argument put forward by Guba and Lincoln (1994) that knowledge is based on multiple interpretations of experience and truth based on particular contexts. In this worldview it becomes important to understand the motives, values and beliefs of individuals. This was necessary in my line of enquiry as the issues I sought to become aware of required some understanding of perspectives and meanings placed on the research questions from differing perspectives. The decision to use, as a local case study, an organisation with which I had some prior involvement, intensified the research issues.

Phase 1 – Participatory evaluation literature review methodology

There is a significant amount of literature that has been produced about different forms of participatory evaluation within the diverse field of development. What is included in this review is mostly material produced after the year 2000 because in that year Estrella et al. (2000) put together 12 case studies that identified who measures change and how competing interests and needs are reflected in indicators; and I also want to ensure that the literature used is relatively current. Exceptions are made for literature produced prior to this that is useful in informing the focus of the research, such as Estrella and Gaventa's (1998) review of literature on participatory monitoring and evaluation.

The literature includes journal articles, books, manuals, reports and policy documents. Of note is *Learning from Change* (Estrella et al., 2000) that is a compilation of international case studies of participatory monitoring and evaluation practice that also provides conceptual and methodological guides on this type of evaluation.

Phase 2 – International case examples methodology

The search for international case examples that involved primary stakeholders in the design and use of outcome indicators was done concurrently with the literature review. The search located very few documented case examples outlining how primary stakeholders have been involved in the design and use of outcome indicators, and evaluation use. Because of the many project areas that the field of development covers including health, environment, education, food security, water, and agriculture, the search for case examples turned out to be a herculean task. There is no one neat parcel online or in libraries where literature on development interventions can be found. Furthermore, there are many organisations involved in the field of development and it was not feasible within the time constraints of this research to go through all their websites in search of appropriate case examples.

Search words used when trawling literature for cases examples included a combination of the terms: i) outcomes, impacts, results, success; ii) indicators, measures; iii) design, development, creation; and iv) development projects, development interventions, development programmes. Although a significant amount of literature matching these search terms was found, most case examples had to be excluded because they did not focus on primary stakeholder involvement in the design of outcome indicators for development

interventions. In some cases, it was unclear if primary stakeholders took part in design of the outcome indicators; the indicators designed were not for measuring progress brought about by a particular development intervention, for instance sustainability indicators designed for a total population's development based on government services; the indicators were designed as early warning signals for e.g. soil erosion. These case studies were not taken into account.

The eight international case examples included in the research are discussed extensively in Chapter Four. Each one is identified using case numbers 01 to 08 in Chapter Four.

Apart from reviewing case examples as a valid means of research for this study, this was also undertaken because the search for a local case study, where primary stakeholders had been, or were going to be, involved in the design of outcome indicators for a development intervention, failed to turn up any examples. As explained in the methodology of the local case study in the next section, a social service organisation about to embark on such a process was identified and used. Rather than rely on data from this case study alone, it was decided to review the case examples as they were based on development principles and practice, and would be valuable for comparison with the methodological considerations identified in the local case study.

The literature review and the case examples resulted in themes being identified that were useful in the development of a rubric. As explained earlier in the chapter, this rubric formed the basis for answering the first research question, using the case examples and the local case study. Themes were drawn from the research data and organised into characteristics of indicators, dynamics of primary stakeholder involvement, and achievements, to answer the other three research questions.

Phase 3 – Local case study methodology

I framed my local enquiry as case study research because, as Simons (2009, p. 443) points out, this is "not a methodological choice but a choice of what is to be studied". Gerring (2004, p. 342) describes case study research as "an intensive study of a single unit for the purpose of understanding a larger class of (similar) units". The fieldwork focused on a single case study that was selected for pragmatic reasons (O'Leary, 2010) that included the timing of my research coinciding with that of the case study organisation, and the uniqueness of the

experience, in that it was the only one I found in New Zealand at the time. The research took on an exploratory nature (Yin, 1984) as I attempted to find answers to questions of a subjective nature, about what appears to be territory that has not been much travelled.

O'Leary (2010) states that interviews, direct observation, participant observation, documents, archive records and physical artefacts are commonly used in case study research. Use of more than one data collection method allows the researcher to corroborate evidence and where there are contradictions, to further explore and understand these (O'Leary, 2010).

The three main sources of case study data are:

- **documents** that included the organisation's annual reports, evaluation reports, assessment sheets, strategic plans, and records of personal communications with organisation members, and the organisation's website;

- **records and notes** compiled during participant observation of steering group discussions, formal and informal discussions amongst organisation members, and focus group activities that were part of the organisation's process; and

- **transcripts** of individual semi-structured interviews with research participants.

Initial contact and preparation for fieldwork

Through a chance conversation about progress with this research, I came to know about what Organisation X, was doing to introduce Results-Based Accountability (further discussed in Chapter Five) as a tool for evaluating, reporting and accountability. While staff members of Organisation X were keen to develop measures for some aspects of the RBA framework, they wanted to involve clients in developing some measures.

I talked to my colleagues and supervisors about the possibility of Organisation X being a case study for this research; Organisation X is a social service organisation that delivers services and programmes for individuals and families to live free of family violence. We reflected on the possible differences between using a development intervention and a social service intervention for this research. One glaring difference between the two was the largely collective nature of the shared goals and planning by a community within a development intervention, whereas for family violence services, setting goals and planning for change as a result of interventions is done at individual or family level. The possible outcomes and

associated indicators for a development intervention would possibly be negotiated and agreed by the group of primary stakeholders. In Organisation X's case, this would not necessarily be the same, as each individual would be working towards meeting their specific individual needs, possibly using different outcome indicators to determine success. What I wanted to examine was the methodological considerations of how the indicators are arrived at by primary and possibly other stakeholders. My concern was not with the intervention itself. I decided that I would approach Organisation X and ask them if they would agree to be a case study for my research.

A total of eight meetings were held with two key informants (KI:1 and KI:2) and other staff of Organisation X. During the meetings I communicated the aim and objectives of my research and they in turn told me about their RBA process and their desire to develop measures using participatory processes.

At the fifth meeting agreement was reached on how I would conduct research on the organisation's process. Terms of reference were drawn up that clearly stated my role and that of the organisation. Organisation X would form a steering group that would guide the process of designing the measures to be used in RBA.

I would, through participation on the steering group, learn about Organisation X and get insight into the development of the process, and, as a researcher who has some experience and knowledge in design of indicators and evaluations, participate in discussion. I would also get to experience the involvement of clients in this process, and have access to documents that would assist with the research. Whatever was produced in this process belonged to Organisation X. I had permission to use my notes and observations from the steering group for my research. Personal notes taken in what could be considered a research journal belonged to me, as did data gathered during my research. My research findings would be shared with the organisation with due respect to anonymity and confidentiality of participants. I was ready to begin my fieldwork but before embarking on fieldwork I had to consider ethics that would guide the process.

Power and ethics

Issues of power dynamics and ethics have implications that researchers have to be attentive to (Ackerly & True, 2008). The conversation about how these would manifest in this research was had with my supervisors, as I prepared to get Massey University ethics approval.

Discussions with my supervisors and senior staff considered whether a low-risk ethics application would be sufficient. Although Organisation X works with women who have experienced and/or continue to experience family violence, the women's experiences in this regard were not the subject of my interest. Rather, my curiosity was in how these women and the staff of Organisation X would operate with autonomy and responsibility to develop outcome indicators that would measure the success of changes in their personal circumstances. Nevertheless, all care had to be taken to ensure that although my research was not to do with family violence, no harm would come to any of the participants. These discussions led to my obtaining in-house ethics approval from the Development Studies Programme in the School of People, Environment and Planning, and subsequent to the completion of a screening questionnaire, submitting a low-risk notification to the Massey University ethics committee.

In preparing for fieldwork with Organisation X, I paid attention to the fact that participants who are clients of Organisation X would possibly feel indebted to the organisation that was or is there for them during such an important part of their lives. I had to ensure they understood that participation in my research was voluntary and that apart from helping me to achieve my academic pursuits, the research would hopefully be useful for Organisation X in considering the processes they were going through. Reflection on how I interacted with members of the steering group, whose members included senior managers and a consultant, was also essential. I wanted other staff to know that I was aware of the power differential that exists between their employers and themselves, and between them and me. I needed to be mindful of my friendship with one of Organisation X's employee's that was declared from the start to both Organisation X and my supervisors. But because of her belief in the integrity of her employer, and my point of contact was with a senior manager, it was considered no conflict of interest.

(Babbie, 2007) points out that researchers share general consensus on the conduct of research, of which the main aspects are voluntary participation, no harm to participants, anonymity and confidentiality, and lack of deception. During fieldwork, I communicated that I was there to observe during the focus groups; participate in discussions and observe dynamics on the steering group; and reiterated the focus of my research during the interviews. Organisation X also explained my involvement to their clients who I would

encounter through their processes. It was in contacting and preparing for the interviews that I spoke to potential participants about:

- voluntary participation and informed consent, explaining that participants who chose to participate could withdraw from the research at any time. An information sheet, given to them before interviews were conducted, further explained the research and the voluntary nature of their participation. Informed consent was obtained by way of signing a consent form (Appendix Two);

- anonymity of research participants. I explained that the information conveyed to me would not be presented as individual stories. Rather it would be organised into themes that would develop from data collected and possibly those that had emerged from the literature review. Organisation X would also not be named in the thesis document;

- confidentiality of research data. None of the data from interviews would be accessible to Organisation X or anyone else, apart from the University if it was required. The transcripts and recordings would be stored on my laptop with password protection and codes for participants;

- potential risk to the participants. As explained, although this research is not about family violence, it involved women who had experienced it. There was a possibility that women would choose to talk about their experiences which could have led to emotional trauma. I had agreement with Organisation X that should this occur I would contact them so that support is made available for the women; and

- taking into account cultural considerations. I had already been told by the organisation that women they worked with were from different cultural backgrounds and embarked on fieldwork cognisant of the need to interact with participants in a manner that was culturally appropriate. I would rely on Organisation X and relationships built with women to learn what this might look like.

Having thought through the dynamics related to power and ethics, I was able to begin fieldwork.

Fieldwork and research methods

The fieldwork took place over a period of about eight weeks between October 2013 and October 2014. During fieldwork I collected data mostly through participant observation. Time constraints meant that data collection stopped before the organisation's participatory process to design outcome indicators was completed. I gathered data on what participants in my research hoped and planned to do as next steps in the process, so that I could get a more complete data set for the research questions I sought to answer. The process of data collection and data analysis was iterative. I analysed data as I collected it to ascertain the need for further data collection. This resulted in participant observation, document analysis and semi-structured interviews being sufficient for answering my research questions.

Document analysis

Use of document analysis was also an iterative process that allowed me to learn about Organisation X's goals and objectives, the work it does, and past evaluations undertaken. I examined documents that were produced by Organisation X for annual reporting, evaluation of pilot services, strategic planning, personal communication and their website. I collected as much documentation as possible, because as Yin (2009) warns, if data collection is incomplete, this can result in biased selectivity. I took heed of Atkinson and Coffey's (2004) advice to examine data within the context they were prepared. In this regard there was awareness on my part that the data collected from documents was not generated to answer my research questions (O'Leary, 2010). An advantage of using document analysis, as pointed out by Yin (2009), included that the documents were always available to read repeatedly, and they contain exact details such as names, events and dates. After my first round of analysis of some of Organisation X's documents, I started participant observation.

Participant observation

I found this to be the most intriguing aspect of my fieldwork, because, as Brewer (2000) points out, participant observation is different from other research methods in that the researcher's experiences during fieldwork are pertinent and legitimate sources of data. This made reflexivity an important practice, discussed later in this section. It was expedient to observe the different aspects of Organisation X's participatory indicator design process rather than to interview people on their participation in the process at a later undetermined date. There were no set timelines of how the design of indicators would be conducted and if I wasn't part of the process, I felt interviews and focus groups may not capture some of the

detail required by my research, such as facilitator skills and expertise in drawing out discussion from clients, nature of participation of clients in focus groups, and how decisions were made.

Participant observation is described by Taylor and Bogdan (1984) as being the social interaction present between the researcher and researched as the latter goes about their day-to-day affairs. In this research, participant observation occurred intermittently over a 13 month period. Taking on the identity of participant-as-observer (Burgess, 1984), I went about forming relationships with stakeholders of Organisation X, most of whom were part of the process I was observing and participating in.

As participant observer I was as the member of a steering group formed by Organisation X to guide their participatory indicatory design process. I felt comfortable taking notes during meetings because, at times, this allowed me to sit apart from the conversations to observe the dynamics of conversation and differing perspectives of those present.

A week before the first steering group meeting, the first focus group was held. Because a senior member of staff was facilitating the process with a client/volunteer, I decided that my observation would be direct rather than as participant. This carried through to the other three focus groups held later. Although my presence had already been explained to participants and there was agreement that I could observe, I introduced myself and explained what my research was about before requesting further verbal agreement to observe their focus group. Each focus group discussed different questions although there were some that were common to all four. These brought out a list of 'Feelings' from the clients that I listed as the clients spoke. The facilitators also wrote a list of Feelings that became the official record.

Apart from the steering group meetings and focus groups I engaged in conversations and meetings, mostly with staff of Organisation X to get greater insight into the participatory indicator design process. Although I collected a significant amount of data during the process of observation, it was important for me that this be complemented by another source. I decided to interview stakeholders.

Semi-structured interviews

A total of eight semi-structured interviews were held with stakeholders of Organisation X (see Appendix Three for interview questions). Four members of staff were

selected because of the roles they had on the steering group and their work within the organisation. I also invited the first person on the four lists I had of clients who participated in the focus groups to interviews. I was however careful not to select a particular client who had begun to access services from Organisation X recently. She had been emotional through most of the focus group organised by Organisation X. I gauged that the risk of "doing harm" was high.

Babbie (2007, p. 320) likens interviews to 'conversations' because of the semi-formal and informal location they take place in, and the potential offered by open-ended questions to allow for exchange to flow naturally. Interviewees chose the interview venues. Of the eight interviews, two took place in Organisation X's office, two in the interviewee's offices, two in the interviewee's homes, one in a restaurant, and one in my own home. All interviews started with about 15 to 20 minutes of chit-chat as we ate food I had brought along. This time helped us to relax and talk about anything the research participants wanted to talk about. Although I explained again that our conversation would not be about their experiences with family violence, but their experiences on the focus groups, this did not stop some from sharing their experiences. I shared my own experiences of family violence with those who wanted to talk about this topic and this seemed to build greater trust in each other and strengthen relationships.

I took into account that the language of evaluation may not be easily understood by participants and that different staff members used different words to mean the same thing. For instance, while some spoke about evaluation, the word analysis was also used, and the terms indicators and measures were used interchangeably. The interview schedule was tested on a friend to whom I had explained the focus groups, their content and the language used by the facilitators. After about 30 minutes of role playing, during which I was the facilitator of the focus group and she, a client of Organisation X, I interviewed her. As advised by Berg (2007), the test run resulted in a more effective interview questionnaire.

Reflexivity

Throughout the fieldwork, what I considered to be my research journal, a folder in which I kept my notes, was useful in many ways. Apart from keeping notes about the fieldwork experience, it became important for reflexive practice.

After the first focus group was held, I read over my notes and reflected on how the focus group had come about and what I observed. After the first steering group meeting I did the same. One of the points I noted after the first steering group meeting was that I felt I had to censure my input into discussion in order not to influence the process. I put these notes in a file with the notes from the first focus group, and those from the preparatory meetings that had been conducted. That file came to be used as a research journal containing my observations and all sorts of diagrams that I formulated to make sense of the various social encounters and my experience of them.

McNay (2000, p. 5) states that reflexivity is "the critical awareness that arises from a self-conscious relation with the other." In my research, this awareness was entrenched in moments of interaction in fieldwork and as I filled out my research journal. Giddens (1991) posits that reflexivity is not only about these temporal aspects, but also about reflection and questioning our epistemologies and the way we see the world, and about the world and what we know of it. A few examples of reflexive practice that were important in my fieldwork are provided.

During the test interview with a friend I found that no matter how I explained the various phases, tasks and roles in participatory evaluation practice, I did not get the answers I sought to a particular question. Afterwards I listened to the interview a few times and later came up with a different way of asking the question. I tried again during the first interview that was with a member of staff of Organisation X. As she tried to answer my question, I realised I still had not framed the question in a way I had hoped would convey its meaning. I felt a sense of frustration with myself and in that moment it occurred to me that I had created the rubric with explanations for each phase (Appendix Four) that had helped me to understand the different phases, tasks and roles in participatory evaluation practice. I had a copy with me and used that to explain what I was talking about. It worked. I reworded the question and used Appendix Four in the remaining interviews. Reflexive practice proved to be effective in the moment.

It was also difficult in all the interviews with clients of Organisation X to elicit responses related solely to the indicator design focus group. There was a tendency by clients to talk about their experiences with Organisation X in general rather than concentrating on the focus groups held. After the first interview, as I made a few notes, I decided to start each question by stating "thinking about the focus group where you talked about your feelings,

past, present and future …" Although this did not altogether stop responses to do with Organisation X, it did lessen these.

I had thought about conducting short telephone interviews with clients who had opted not to take part in Organisation X's focus group discussions. As I sat to prepare an interview schedule, I tried to put myself in these women's shoes. I realised then that they had opted not to participate for varying reasons. By calling them I might make them feel unnecessarily guilty for not contributing to the work of an organisation that supports them. I was reminded of Cooke and Kothari's (2001) warnings about the tyranny of participation in which they contend that participation is not always desirable and that participatory methods could get in the way of using others that are potentially more productive in attaining the aspirations sought through participation. In some way I had started to romanticise the non-attendance of women to focus groups as a way of asserting their choice to not take part. Reflexive practice in this instance led to the realisation that the reasons for not participating could have been due to logistical issues. I was willing to accept that in research there are times when you leave the field with questions that remain unanswered.

Data analysis and coding

During the data coding and analysis stage I was guided by O'Leary's (2010) advice that entails organising raw data into categories, coding data, and then analysing it. Data analysis loosely followed template analysis and analytic deduction. N. King (2004) explains that template analysis is a group of techniques used to organise and analyse textual data thematically. It can be used from a contextual constructivist position that recognises the relationship between data and findings, and the context within which they are produced (Madill, Jordan, & Shirley, 2000). Reliability of coding becomes less important than reflexivity of the researcher, and representation of the data using rich descriptions to convey the meaning of the text. The analysis of data also borrowed from analytic deduction that P. Johnson (2004) states is a process in which data is organised into components that are scrutinized to form patterns and relationships resulting from the literature. I began analysis of data by immersing myself in it.

Once all the data were collected, I had to label and organise it, making it ready for analysis. Labels were given to each of the participants and other data sources as follows:

- Meetings were held with two key informants at the start of the process. These key informants are labelled KI:1 and KI:2;

- Four focus groups held by Organisation X. Each is labelled FG1-4 and participants labelled FG1:1 - 4, FG2:1 - 4, etc.;

- Four interviews held with clients. Participants are labelled C1-4;

- Staff who took part in the research included senior management of the organisation and operational staff. Research participants were interviewed because of the role they had played during the participatory indicator design process. A consultant who volunteers her time with the organisation was also interviewed. Three staff and the consultant are labelled S1 – S4; and

- Staff and steering group members also participated in a workshop to discuss frameworks. All those who participated were labelled WP1 – WP8.

Transcribing of data and making sense of notes and diagrams from meetings provided a vehicle through which I became immersed in the data. By choosing to interpret the data on my own I was able to decipher recordings that had been affected by the sound of birds and wind during an interview carried out in one interviewee's garden, and the sound of construction equipment in another interview. Bailey (2008) suggests that transcribing be done by someone familiar with the context because of the poor quality recordings described, and to facilitate immersion in the data. Reflexivity was important during this process. Before transcribing the eight interviews, I listened to them to remember what each interviewee was conveying through body language, their words, and often laughter. I transcribed and gave the documents to each of the interviewees to give them a chance to review what they had shared with me. Upon return, I re-read the text and started to draw out themes. I listened to some of the interviews again, this time focussing on how my positionality and motivation affected analysis, and where I might have chosen to place emphasis. I found that this allowed me to see the text differently and added another dimension to my findings.

The data from the interviews, participant observation, and document analysis is organised into themes, namely Phases of involvement in participatory evaluation practice; Characteristics of indicators; Dynamics of involvement and lessons learned; and Achievements. These themes were further broken down to reflect the sub-headings in the

review of the eight case examples. This coding of data formed the template for analysis. During analysis I was also able to link emerging themes to the literature review. This process helped to organise the findings and discussions sections.

Chapter summary

This chapter started by outlining the research questions and providing rationale for use of a constructivist approach. It outlined the three phases of this research namely the literature review, the review of international case examples and the local study that are presented in Chapters Two, Four and Five respectively. The influences of constructivism in finding and interpreting meaning are also put forward. It outlined the methodology for each of the research phases and how the literature review was significant for identifying key themes to follow up on in the review of case examples and the local case study.

The fieldwork experiences and research methods were discussed. A brief explanation of the data collection methods, namely document analysis, participant observation and semi-structured interviews, and their uses in fieldwork were presented. Of note is the use of a research journal during fieldwork. A useful process that emerged from its use was that of reflexivity in the moment, and when filling out my research journal. The importance of being able to flexibly adapt my research method and tools in order to find new and different ways of relating to research participants was evident.

The chapter draws to an end with an outline of how data were analysed by loosely following template analysis and analytic deduction. The effectiveness of transcribing data as a technique to immerse myself in it was also presented.

Chapter Four: Review of case examples

Introduction

This chapter reviews literature reporting on international case examples in which primary stakeholders have been involved in the design and use of outcome indicators, and evaluation use. The review aims to provide an understanding of the phases that primary stakeholders have been involved in; the design of outcome indicators and evaluation practice; the characteristics of indicators they design; the dynamics of involving primary stakeholders in such processes; and what is achieved through this involvement.

This review of case examples builds on the theoretical and practical aspects of participatory evaluation, and indicator design specifically, provided in Chapter Two. It specifically seeks to identify the methodological aspects of participatory indicator design as part of the entire evaluation process. In doing so, it is noted that the case examples are drawn from available literature that was not written for this particular study. Care is therefore taken in interpretation of the data available and ensuring that absence of evidence is not interpreted to mean evidence of absence.

Case examples reviewed

This section gives an overview of the case examples reviewed, including a summary of the cases examples in Table 4.1. It then moves on to analysis of primary stakeholder involvement in the four phases of participatory evaluation practice using the rubric designed in Chapter Two. Figure 4.1 shows the evaluation process.

As explained in Chapter Three, the search for international case examples that matched the criteria of this research resulted in only a few being identified as suitable. The international development case examples are from a variety of fields including sport for development (Holte-McKenzie et al., 2006), natural resource management (Hillyer & Purohit, 2007), conservation (McDuff, 2001), joint management (Izurieta et al., 2011), agriculture (Parkinson, 2009), gender and development (Scheyvens, 2007), and research and development (Njuki et al., 2008; Vernooy et al., 2006). Table 4.1 offers a summary of these case examples.

Table 4.1: Summary of case examples reviewed

Case Number and Country and Organisation/Research Team	Reasons for involving primary stakeholders
Case 01 **Kenya: Wildlife Clubs of Kenya**, a youth organisation, is the largest grassroots conservation organisation in Africa.	To build the capacity of the organisation to conduct **participatory evaluation**. The organisation wanted to increase use of evaluation findings by ensure ongoing collection of data that is responsive to community needs and for conservation. (McDuff, 2001)
Case 02 **Kenya: Moving the Goalposts Kilifi** is a Sport for Development organisation whose aim is to work with girls and young women to harness their life skills and capacity to contribute to local community development.	To develop a **participatory monitoring and evaluation** strategy for the organisation. (Holte-McKenzie et al., 2006)
Case 03 **China: Guizhou Academy of Agricultural Sciences** is a research and development organisation that has research activities and interventions on the use of water, farmland, grassland, forests and wasteland.	To strengthen the organisations' development research through capacity building in use of **participatory monitoring and evaluation.** (Vernooy et al., 2006)
Case 04 **India:** A multi-agency multi-disciplinary research team working with **self-help groups** in peri-urban areas to increase capacity of communities to adapt natural resource management and livelihoods strategies to their changed environments.	To design indicators for their self-help groups' objectives and to develop **participatory monitoring and evaluation** plans. (Hillyer & Purohit, 2007)
Case 05 **Tanzania: Community Organisation for Research and Development**, an organisation that works with pastoralist peoples in Northern Tanzania.	To develop a **monitoring and evaluation** plan that ensured "women from the women's groups play the key role in identifying appropriate indicators, and they should also play a strong role in monitoring progress." (Scheyvens, 2007)
Case 06 **Malawi:** A **team of researchers** who used a research for development framework that seeks to strengthen the capacity of farmers to have productive connections to markets, to create lucrative agro-enterprises, and to use technologies in a manner that enhances food security and sustainable enterprise.	To apply indicators developed by six communities for **participatory monitoring and evaluation** purposes. The team wanted to aggregate the indicators into three categories: livelihood, human capital and empowerment, and social capital, and to assess community perceptions of the project's achievements. (Njuki et al., 2008)
Case 07 **Uganda:** The **National Agricultural Advisory Services** seeks to help farmers to access information, technology and knowledge for profitable agricultural production.	To introduce **participatory monitoring and evaluation** to enhance organisational learning and programme management, and to involve particularly poor and marginalised farmers in the process as a means of empowerment. (Parkinson, 2009)
Case 08 **Australia: Government Parks and Wildlife Services of the Northern Territory of Australia**, the Lands Council that represents traditional Aboriginal land owners, and the Aboriginal people of Flora River Nature Park, Daminmin National Park, Watarrka National Park and East MacDonnell Rangers Park. Focus is on Flora River Nature Park for which there is a considerably large amount of data available.	To develop joint management plans for four parks using **participatory monitoring and evaluation** and joint management principles, and to then consider "joint management effectiveness through the participatory development of criteria and indicators" and to see how they performed in an evaluation. (Izurieta, Petheram, Stacey, & Garnett, 2013; Izurieta et al., 2011; Stacey, Izurieta, & Garnett, 2013)

Source: Author.

In evaluation research, frameworks and models provide structure and guidance on the manner in which the process will be conducted. These frameworks and models are often adapted to suit a particular context.

Approaches in evaluation research provide different conceptual ways of designing and doing evaluations. Those used in the case examples reviewed include participatory monitoring and evaluation (Cases 03, 04 and 07), Participatory Rural Appraisal (Case 06), Participatory Action Research (Case 08), Participatory Learning and Appraisal (Case 02), monitoring and evaluation (Case 05), and a Training of Trainer model created to specifically cater for wildlife conservation projects (Case 01),

Varying frameworks and models were used in the case examples reviewed. For the National Agricultural Advisory Service programme, the Logical Framework Approach was used and the logframe was the starting point for forming the data collection tool (Case 07). The joint management process for the protected areas in Northern Territory of Australia (Case 08) utilised two frameworks: the World Commission on Protected Areas framework for assessing management effectiveness of protected areas, and the Capital Assets Assessment. A Training of Trainers model was specifically developed by the project stakeholders and facilitator for the Wildlife Clubs of Kenya participatory evaluation (Case 01). The indicators used by the research team in Case 06 were placed within the Sustainable Livelihoods Framework.

Phases of involvement in participatory evaluation practice

This section seeks to answer question one of the research: At what phases in participatory evaluation practice are primary stakeholders involved?

While reviewing the case examples, four phases of indicator design were identified: 1) Planning; 2) Indicator design and use, 3) Data handling, and 4) Evaluation use. The proposed rubric introduced in Chapter Two to frame classification of the eight case examples, was designed using guidance on participatory evaluation and indicator design offered by Reed et al. (2006), Guijt (2000), Estrella and Gaventa (1998), and Aubel (1999). Common themes from this literature were drawn out that allowed for the rubric, created in Chapter Two, that offers a broader typology against which to classify involvement of primary stakeholder participation to be developed. Figure 4.2 depicts a linear process of the four phases. The process of involving primary stakeholders is however not linear, and is often iterative (Hillyer & Purohit, 2007; Holte-McKenzie et al., 2006) to take into account ongoing developments that may or may not have been taken into account during planning. Appendix Four provides the detail of the phases, tasks and roles.

Figure 4.1: Phases of involvement in participatory evaluation practice

Phase 1 - Planning

Decision to design indicators/ Decision to evaluate

Guide research/evaluation process

Selection research/evaluation framework

Needs assessment

Capacity building

Phase 2 – Indicator Design and/or

Data collection

Design of Indicators

Selection of indicators

Aggregation indicators

Indicator Use (Evaluation)

Phase 3 – Data Handling

Data analysis

Report writing

Information dissemination

Phase 4 – Evaluation Use

Implementation of evaluation findings

Source: Author.

Phase 1 - Planning

The planning phase involves the decision to design outcome indicators or evaluate; selecting people who will form the evaluation or research team; selecting a framework to be used; a needs assessment; and capacity building of those involved in the process.

Often the decision to design outcome indicators and/or evaluation plans, or to conduct evaluations, is made by donor or funding agencies (Cases 04 and 05) and/or the organisation implementing the intervention (Cases 01, 02, 03 06, 07 and 08). It would appear from the eight case examples that primary stakeholders are not often asked at what level they would like to participate in these processes. Once the decision to design indicators in a participatory way has been made the involvement of primary stakeholders is considered.

In cases 02, 03 and 08 there is evidence that primary stakeholders are part of the research or evaluation team to help make decisions about how outcome indicators will be designed and/or evaluations conducted. Few of the case examples reviewed made reference to this.

In Cases 01, 02, 03 and 08 there is mention of training of primary stakeholders in aspects of evaluation processes. In Case 01 farmers were trained in the collection of data at local and regional level (McDuff, 2001). A small number of primary stakeholders were trained to train others in mostly understanding what evaluation is, design of indicators and data collection in Case 02. Together with staff of a research organisation, farmers, who were primary stakeholders in Case 03, participated as part of a research team. A capacity building

process on evaluation was conducted for the multi-stakeholder team in Case 07. In Case 08 the facilitator worked with members of staff and primary stakeholders in data collection and other aspects of evaluation (Izurieta et al., 2011).

Phase 2 - Indicator design and use

The design and use phase involves the design of outcome indicators and/or the implementation of evaluation. The design of outcome indicators is often considered the most crucial aspect in participatory monitoring and evaluation (Guijt, 2000). Because this literature review focused on participation in the design of outcome indicators, most of the case examples demonstrated high levels of participation in this aspect. In Case 02 the design of outcome indicators was done by primary stakeholders on their own. In Cases 01, 04 and 08 primary stakeholder participation occurred alongside other stakeholders, and in at least two cases (05 and 07) it appears primary stakeholders had input into outcome indicators proposed by staff of the organisation. The extent of participation in the other two case examples is not clear.

It seems that in a few case examples the indicators designed by and with primary stakeholders were tested. After testing the indicators designed with self-help groups in Case 04, it was found that some were not appropriate. Some self-help groups adjusted their plans and indicators, as the testing period gave them better understanding of what indicators would provide useful information (Hillyer & Purohit, 2007). For Case 01 the process started with the design of evaluation questions and indicators. These were then tested on one of the local school programmes (McDuff, 2001).

Scheyvens (2007, pp. 9-12) speaks of "the idiosyncrasies of indicators" and how during design phase indicators may seem to be appropriate, but after they have been tested or used in an evaluation, it is found that they are not suitable and need to be changed in order to collect relevant, useful data. It appears in Case 05 that indicators were designed by staff of Community Organisation for Research and Development with the researcher, with possible input from primary stakeholders. It would also appear that discussions with stakeholders resulted in more appropriate indicators being designed.

Depending on the type of indicators designed, it may also be possible to assess them against existing indicators or frameworks as the research team in Case 06 did. (Njuki et al., 2008) used the Sustainable Livelihoods Framework that focuses on natural, human, social,

physical and financial assets to assess the indicators designed by local farmers. This was also done in Case 08 where their indicators were assessed against the Sustainable Livelihoods Framework and the World Commission on Protected Areas management effectiveness framework that takes into account context, planning, inputs, process, outputs, and outcomes (Stacey et al., 2013). Although Izurieta et al. (2011) do not make specific reference to how the indicators fared within the two frameworks, based on two tables provided in the article it would appear that indicators aligned with those in both frameworks.

Three of the eight case examples mentioned use of indicators in evaluation. In Case 01, (McDuff, 2001) worked with a multi-stakeholder group of the Wildlife Clubs of Kenya to build their capacity on participatory evaluation. This resulted in an evaluation to test the indicators and involved primary stakeholders as evaluators.

Indicators developed in Case 08 were also used in an evaluation in which primary stakeholders were involved. Traditional land owners, the primary stakeholders in this case, preferred to participate as interpreters rather than conduct interviews "because they felt more confident in this role ... particularly given that this activity was new to them and people were generally more comfortable speaking in their aboriginal language" (Stacey et al., 2013, Data Collection Section).

Case 06 utilised the indicators developed by farmer communities for evaluation purposes. It is unclear whether primary stakeholders were involved in the aggregation of indicators, and the evaluation that made use of these indicators. In Case 02, available literature on evaluations of the organisation's programmes do not utilise indicators that were designed by the girls and young women, even though in at least one of the two instances evaluators were aware of the indicators (Woodcock, Cronin, & Forde, 2012).

Phase 3 - Data handling

This phase involves data analysis, reporting and information dissemination. While there is evidence that data analysis in all cases occurred, it is often not made clear how primary stakeholders were involved. Only in Case 02 is there mention of primary stakeholders being involved in analysis of the data collected and reporting on it. In the other cases it appears this was done by the research/evaluation team, often made up of staff and donors.

Phase 4 - Evaluation use

When primary stakeholders participate in the processes of designing and using outcome indicators and evaluations, utilisation of findings increases (Patton, 2008). Case 08 alludes to ongoing involvement of primary stakeholders in evaluation processes because this is done as part of a joint management plan (Izurieta et al., 2011). The evaluation conducted in Case 03 led to the development of a self-monitoring tool in the form of a booklet in which householders were to design their own indicators for monitoring water management (Vernooy et al., 2006). This resulted in the "enhanced villagers' capacity to identify problems in water management and to find effective solutions for the problems" (Vernooy et al., 2006, p. 406).

Based on the review of the eight case examples, primary stakeholders are often not involved in all aspects of participatory evaluation practice. The analysis of involvement of primary stakeholders in the four phases is further discussed in Chapter Six.

Characteristics of indicators

The second question of this research seeks to identify the characteristics of the outcome indicators designed by and with primary stakeholders. While reviewing the eight case examples, four categories emerged to form the basis of the key characteristics of indicators designed by primary stakeholders. The four categories are to do with indicator: 1) links to goals and objectives; 2) negotiations, 3) nature; and 4) quality. The categories are used to discuss the indicators.

Linked to goals

Cases 02, 05 and 08 pointed to the importance of indicators being linked to the goals and objectives of an intervention. Holte-McKenzie et al. (2006) and Izurieta et al. (2011) state that when goals and objectives are the starting point, in the design of indicators, the links are made clear to the evaluator. Other case examples alluded to indicators designed to meet the objectives of their particular development interventions.

Negotiated

Competing interests about what should be measured in an evaluation and the manner this should be done can result in differing views on monitoring and evaluation and the indicators to be used. Because indicators reflect the interests of different stakeholder groups

and/or different interests within one stakeholder group, it is important that the final indicators are negotiated.

In Case 01 debate and discussion on the indicators to be used occurred amongst different stakeholder groups to ensure they were relevant for the programme. The negotiations in Case 08 took place due to competing interests of the stakeholder groups. Facilitators were careful to ensure that criteria and indicators for measuring joint management performance reflected the concerns of all partners.

In Case 04 it seems that there were negotiations between the research team and primary stakeholders in the design and selection of indicators as there are reports of substituting indicators designed by primary stakeholders

Nature of indicators

Each case presented different information on the nature of the indicators designed by primary stakeholders. The case examples commented mostly on indicators being quantitative and/or qualitative; designed based on cultural understanding and lived experience; visual to aid understanding and simple to use.

Culturally appropriate

In Cases 02 and 06 it was found that indicators designed by primary stakeholders often reflect the realities and lived experiences of communities within which interventions are carried out. In relation to Case 02, Holte-McKenzie et al. (2006, p. 372) state that "the final indicators chosen for leadership emphasised the behaviour and comportment of the players ... rather than characteristics the principal researcher would have chosen, reflecting more her own cultural background than the one the research was conducted in". Coming from two different cultural perspectives, it was rare for shared objectives and indicators to be identified and agreed by the staff of Flora River Nature Park and the Aboriginal land owners in Case 08.

Simple

Because multiple stakeholders take part in participatory evaluation, it is important that they are able to understand how the indicators chosen will measure change. Cases 02, 04 and 08 encourage the design of simple indicators. Accordingly Izurieta et al. (2011) posit that indicators developed by primary stakeholders are likely to be easier to use in participatory evaluations that involve primary stakeholders, than those developed by experts because they

are developed within a particular local context. These simple indicators increase the chance for programmes to meet the needs of primary stakeholders and there is likely to be a decrease in indicators that reflect assumptions held by programme designers (Izurieta et al., 2011). Designing of indicators was associated with selection of methods that are cost effective and able to be utilised within existing opportunities like soccer tournaments, from the Case 02 evaluation plan.

Hillyer and Purohit (2007) stated that visual aids in the design of indicators and evaluations are valuable not only for assisting with understanding of concepts, but also for those who are illiterate.

Qualitative and quantitative

There is an argument for having both qualitative and quantitative indicators for interventions. In all the case examples there is evidence that the indicators designed include both quantitative and qualitative ones. In Case 05, although the team designing indicators realised from the outset that qualitative indicators are useful, the decision was made to include quantitative ones. They found that including quantitative indicators was not always appropriate because the manner in which information was collected made it difficult to measure change. Scheyvens (2007) concludes that anecdotal and qualitative data, to record change in the self-confidence and skills of women in Case 05, were more appropriate than the quantitative measures chosen in the first instance.

Location and context specific

The belief that indicators for interventions need to be locally relevant was a factor in considering inclusion of primary stakeholders in design of indicators and evaluation (Holte-McKenzie et al., 2006; Njuki et al., 2008; Scheyvens, 2007). In Case 02 it was found that local relevance of indicators is important because outcomes are affected by context. The organisation was more concerned with indicators that were appropriate for the girls and young women than they were in being able to replicate these to other development interventions (Holte-McKenzie et al., 2006). Commenting on the indicators developed by the girls and young women, the Moving the Goalposts Kilifi program coordinator asserted that the girls had "really grasped leadership in the context in which they live" (Holte-McKenzie et al., 2006, p. 372) reinforcing the view that outcome indicators created by primary stakeholders for a particular intervention can be location-specific.

There are ways of ensuring that indicators can be used across multiple sites with similar development interventions. To achieve this, the facilitator in Case 01 worked with stakeholders from multiple sites to develop indicators and test them on their wildlife programmes (McDuff, 2001). In Case 06 the team of researchers aggregated indicators devised by local farmers and compared these across communities. They concluded that the indicators developed by the primary stakeholders needed to be complemented with more quantitative ones, and grounded in more theoretical frameworks for comparison with other generic indicators, so that community progress is not assessed from a narrow perspective (Njuki et al., 2008).

Quality of indicators

The quality of indicators in the case examples is mostly judged by their ability to represent the reality within which they are constructed. In Case 08, Stacey et al. (2013, Conclusion Section) conclude that the indicators "identified and measured ... should be seen as a statement of the current status". The same was found in Case 02. Most indicators designed by the girls and young women were related to football and Holte-McKenzie et al. (2006, p. 368) note that it was not easy to "draw out other experiences related to life off the pitch" . The article does not state how and if they overcame this issue.

The research team working with self-help groups in Case 04 found the value of indicators is affected by the process used to identify them. In their experience a narrow focus on participation during the design of indicators alone was not sufficient. They argue that quality was enhanced when the self-help groups were involved in other parts of the process, including deciding how they would be used, by who and how often (Hillyer & Purohit, 2007, p. 16).

Dynamics of involvement and lessons learned

The third research question seeks to understand the dynamics of involving primary stakeholders in the design and use of outcome indicators, and evaluation use, and the lessons learned. The eight case examples provided insight into the dynamics of such involvement. The review found five issues of particular relevance: 1) the role and expertise of facilitators; 2) influence of culture on the process and the indicators designed; 3) participation – including power, manner of participation, and how participation was enhanced; 4) the amount of time it takes to complete the process; and 5) associated costs.

Facilitation

During the design of outcome indicators and evaluations provided in the case examples reviewed it is made clear that the skills of the facilitator's affect the outcome of the process. Table 4.2 presents data on which stakeholders participated as facilitators in evaluation processes.

Table 4.2: Facilitation in case examples reviewed

	FACILITATION		
	Primary Stakeholders	Staff	Facilitator
CASE 01			X
CASE 02	X		X
CASE 03	X	X	X
CASE 04		X	X
CASE 05		X	X
CASE 06			X
CASE 07		X	
CASE 08			X

Source: Author.

Of the eight case examples, primary stakeholders participated as facilitators or evaluators alongside other stakeholders in two cases; staff of the organisation or institution participated in four; while facilitators participated in seven. The different stakeholders had varying degrees of knowledge on the subject area for which indicators were being designed and differing skills and expertise on facilitating such processes.

It would appear that in Case 01 there was a clear process of how the evaluation planning would occur and that the facilitator was able to see this through without any adverse issues being reported (McDuff, 2001). The experience started later than planned in Case 08 because neither joint management partners, i.e. staff of Flora River Nature Park and the Aboriginal land owners, were keen to embark on the process of participatory monitoring and evaluation without adequate experience and an external facilitator. The researchers were able to "play a key role in minimizing biases by enabling equal space for participation and contribution (Izurieta et al., 2011, Conclusion Section, para. 9). Having facilitators provided extra human capacity and played a key role in promoting closer working relationships between the partners. The experiences described in these two case was not however common to all the case examples.

In Case 02 the girls and young women were trained to facilitate processes. Holte-McKenzie et al. (2006, pp. 372-373) concluded that "the attitude and skill level of the research

team may have influenced how comfortable participants felt and to what extent they participated" because "it is possible that the words or tone used by the team when conducting FGDs did not put the participants at ease".

Lack of adequate skills and expertise impacted on Case 04. The research team reportedly had varying degrees of expertise on participatory processes. The conclusion reached was that "the team's own lack of clarity frustrated attempts to facilitate the process" (Hillyer & Purohit, 2007, p. 16). One of the difficulties encountered by the team was lack of knowledge on the differences between "objectives and indicators ... indicators and methods ... methods and measures" (Hillyer & Purohit, 2007, p. 16). The research team stated that because the process was said to be unenjoyable and without reward, it was unlikely the process would continue beyond the life of the project. Hillyer and Purohit (2007) found, however, that the process resulted in the facilitators gaining new skills such as knowing when to give their input and intervene without dominating; and recognising the importance of participation in monitoring and evaluation.

Culture

As stated earlier, the differing cultures of the staff and the Aboriginal land owners in Case 08 meant that apart from seeing the process from two different perspectives, they had differing goals. The Aboriginal perspective on joint management is culturally different from that of the government (Izurieta et al., 2011). For Aboriginals the joint management process is connected to community development in which their desire to control traditional lands is a realisation of stronger cultural identity and self-determination, while the government views joint management as being about attaining conservation goals. The Aboriginal land owners were able to share their culture with partners and the public through the process (Izurieta et al., 2011).

In Case 02 it was found that the girls and young women took part in processes in the manner they understood and interpreted participation. On further inquiry it was established that rather than being encouraged in their schools, participation is restricted. This is a possible explanation for the limited way the girls and young women participated during the process.

Participation

In most of the case examples reviewed participation was limited to those who were directly affected by the intervention or process. One issue that has a bearing on the manner of participation is power dynamics between and within stakeholder groups. Findings on power dynamics and the manner of participation are presented.

Power dynamics

Unequal power dynamics exist for reasons that include cultural practices and social mores, age, skills and educational levels of participants (Holte-McKenzie et al., 2006; Izurieta et al., 2011), the attitudes of those involved (Holte-McKenzie et al., 2006), and the role and skills of the facilitator (Hillyer & Purohit, 2007). These dynamics, which are often intertwined, can be a barrier to participation in evaluations. There are however ways in which evaluation practice can dilute the power of the facilitator and the dominant voices amongst stakeholders.

Dynamics amongst stakeholder groups - Forming relationships amongst stakeholder groups can contribute to dilution of power. Case 08 shows how this was achieved by having external facilitators and ensuring that staff and the Aboriginal land owners spent time together informally. This brought about greater understanding by staff of the cultural practices, norms and values of the Aboriginal people, and allowed for flexibility in the manner in which design of indicators was carried out. It also allowed for strengthening of personal and working relationships (Izurieta et al., 2011).

The importance of relationships between facilitator and stakeholders is also seen in Case 07. Parkinson (2009) pointed out that there were numerous assumptions made by the different stakeholders on the need for participatory monitoring and evaluation and who would benefit from it. Power imbalances existed between staff and farmers and amongst the famers themselves (Parkinson, 2009, p. 235). Farmers felt that participatory monitoring and evaluation had been introduced by programme managers for the benefit of the organisation because it was thrust upon them suddenly without prior discussions. The article does not state if and how the organisation addressed these imbalances.

Dynamics within stakeholder groups - Although primary stakeholders of different ages, social standing and educational attainment may be invited to take part in the design of outcome indicators together, in Cases 02 and 08 it is reported that younger participants tend to defer to older participants for direction in their input.

Coupled with age, social standing tended to cause unequal participation amongst the Aboriginal people of Flora River Nature Park in Case 08. Izurieta et al. (2011) noted that older people participated more than the younger ones because they had authority within their community and were formally or informally appointed as spokespersons.

For the girls and young women in Case 02 it was found that younger girls concurred with what older girls had to say. Educational attainment also contributed to this as younger girls seemed to ask older more educated girls what indicators they would vote for (Holte-McKenzie et al., 2006). The experience in Case 07 was that although it was known that there was high level of illiteracy amongst farmers, the tools for collecting data appear to have disregarded this, as most farmers found it difficult to fill out the forms (Parkinson, 2009). The illiterate farmers felt they "were likely to be marginalised from programme benefits at the expense of more literate groups" and the forms "created a perceived division between poorly educated and more educated farmers … within groups and between groups, since some groups lacked any member who was confident in dealing with the forms" (Parkinson, 2009, p. 235).

Manner of participation

The manner in which primary stakeholders participated varied. Some participation can be described as "active" in that participants contributed their views eagerly during processes, while others were "passive" in that although they turned up, they contributed little to the process even though they may have been interested in what was going on.

In Case 08, participation was said to be passive because although most of them were interested in knowing what was going on, they did not have any decision-making powers over the land and therefore said little. According to Izurieta et al. (2011) older adults who had knowledge of Flora River Nature Park played a bigger role than others in vocalising their views. Participation was also affected by lack of adequate literacy and numeracy skills, limiting participation to only the visual and oral aspects of the process (Izurieta et al., 2011).

Enhancing quality of participation

Because stakeholders sometimes use different languages to communicate, challenges arise during evaluation processes. The Aboriginal land owners, who had limited use of English, tended to translate for each other during meetings (Izurieta et al., 2011). Lands Council staff enhanced quality of participation by having pre-meetings with Aboriginal land

owners to discuss the content of the proceeding session, as did an anthropologist who assisted with Aboriginal engagement. The evaluation team formed in Case 02 included a local translator who was the same age as some of the young women in the intervention to ensure language was not a barrier (Holte-McKenzie et al., 2006; Izurieta et al., 2011).

The use of visual aids to facilitate the process is encouraged especially where some of the participants are illiterate. In Case 04 the research team worked with self-help groups to develop their own visual indicators.

In Case 02 the girls and young women formed a planning group that worked closely with the evaluation team. This allowed for greater participation and understanding of evaluation processes (Holte-McKenzie et al., 2006).

Length of time

Working with the partners of Flora River was a time consuming process. Flexibility was required in Case 08 when the Aboriginal people had to take part in events such as funerals, as this disrupted the process (Izurieta et al., 2011). As a result, the process took longer than the facilitators and staff thought it would.

The process of designing indicators in Case 04 was time consuming and described as "arduous" (Hillyer & Purohit, 2007, p. 16). At one point this led to a feeling of lack of enjoyment, reward, and useful learning. On reflection the team reviewed their process so that the groups worked on one indicator at a time i.e. design, decisions of how it will be used, etc., rather than trying to design all the indicators, then looked at measures for all of them, etc. (Hillyer & Purohit, 2007, p. 17). This reportedly helped with understanding, especially with groups who had already put a lot of time into trying to develop their processes.

Cost

The cost of participatory monitoring and evaluation and working with primary stakeholders to design indicators was addressed in only Case 08. Izurieta et al. (2013) worked out the cost of participatory monitoring and evaluation at four parks, including Case 08, and found the cost was almost the same for all sites and constituted 1.3% of the total cost of joint management (Izurieta et al., 2013). Izurieta et al. (2013) stated that the most expensive aspects of participatory monitoring and evaluation were the preparation (taking up half the cost because it involved training teams in participatory monitoring and evaluation), and

validation phases. The single most expensive aspect was wages, including facilitation costs. Other costs of participatory monitoring and evaluation were at times absorbed into the operating budget of joint management, as the two were integrated. Izurieta et al. (2013) therefore show that it is possible to combine participatory monitoring and evaluation aspects with internal processes to keep costs lower than the 5-10 percent of the budget of any biodiversity management, as recommended by Lindenmayer et al. (2012).

Achievements

The last research question seeks to find out what is achieved when primary stakeholders are involved in the design and use of outcome indicators, and evaluation use. Positive outcomes are reported from the case examples reviewed. Most of the cases were said to have achieved empowerment of primary stakeholders, designing appropriate indicators and plans. This section addresses achievements realised in the case examples in relation to these aspects.

Empowerment

Vernooy et al. (2006, p. 401) state that participatory monitoring and evaluation is potentially empowering "because it puts local people in charge; helps to develop their skills; shows that their views count; and provides an opportunity for joint learning." This section therefore looks at whether the case examples reviewed achieved empowerment in this regard.

Were primary stakeholders in charge?

To some extent primary stakeholders were involved in phases and roles in which they had control of evaluation processes. Being on the evaluation team and the planning group in Case 02, and based on the experience reported, the girls and young women were responsible for decisions on data collection tools, designing and selecting indicators, data collection, testing of indicators and reporting. In Case 08 the process provided a "baseline from which the level of equity in representation and opportunities for input to planning and decision making … as part of an equitable partnership" (Izurieta et al., 2011, Discussion Section, para. 7). In these two cases it was clear that primary stakeholders were in charge of some aspects of evaluation.

Were the skills of primary stakeholders developed?

The skills of primary stakeholders were developed in a variety of ways in most of the case examples reviewed. In Cases 01, 02, 03, 04, 06 and 08 there is evidence of capacity building in evaluation processes beyond indicator design. Local stakeholders in cases 05 and 07 benefitted from training in at least designing outcome indicators.

Did the views of local people count?

The ways in which the views of local stakeholders were taken into account varied. This included local perspectives on: the indicators themselves in all case examples, methods used to measure change (Cases 02 and 04); interpretation of meaning given to aspects such as joint management (Case 08); enhancing participation (Case 01 and 02); and giving input on the "reporting format based on the … logical framework" (Parkinson, 2009, p. 231).

Was there opportunity for joint learning?

In the case examples where more than one stakeholder group was involved there appears to have been joint learning. The case examples highlighted different points of learning including incorporating the traditions of an organisation such as competitions into evaluation processes (Case 01); applying Participatory Learning and Appraisal for a sports-for-development project (Case 02); improvements of water management systems in Guizhou, China (Case 03); the best methods and measures to use for self-help groups in India (Case 04); working out whether quantitative and/or qualitative indicators were appropriate measures (Case 05); and conducting participatory monitoring and evaluation as part of joint management (Case 08). Table 4.3 presents aspects of empowerment described in the eight case examples. It is based on Vernooy et al.'s (2006) criteria for empowerment

Table 4.3: Aspects of empowerment noted in case examples reviewed

ASPECTS OF EMPOWERMENT	CASE 01	CASE 02	CASE 03	CASE 04	CASE 05	CASE 06	CASE 07	CASE 08
Control		X	X					X
Skills development	X	X	X	X				X
Local views	X	X	X	X	X	X	X	X
Joint learning	X	X	X	X				X

Source: Author.

It would appear that inclusion of primary stakeholders to inform local views in the development of monitoring and evaluation plans was highly desirable in all the case examples.

Three case examples address the issue of control through their account of how primary stakeholders were involved in processes where they were able to have influence in decision-making. Joint learning and skills development of primary stakeholders was apparent in five of the case examples.

Appropriate plans and indicators

The indicators developed in Case 01 by the multi-stakeholder group were such that Wildlife Clubs of Kenya was able to use them and associated tools to create a baseline for monitoring and evaluation, partly because the primary and other stakeholders linked these to competitions, already embedded in the organisation's culture. The process also highlighted the need for information management systems within the Wildlife Clubs of Kenya (McDuff, 2001).

In Case 03 it was noted that involving farmers resulted in better quality plans and appropriate indicators, because the original plan and tools presented by the research team to the farmers were modified to make them "farmer-friendly" (Vernooy et al., 2006, p. 405).

Self-help groups in India reportedly achieved some effective participatory monitoring and evaluation processes in Case 04. According to Hillyer and Purohit (2007), the methods identified with the self-help groups proved more valuable that those developed by the research team.

The plan developed for Case 08 was the result of joint agreement on the future directions of Flora River Nature Park's day-to-day management, including the identification of management outcomes against which monitoring and evaluation could be carried out (Izurieta et al., 2011). The plan reflected the aspirations of the Aboriginal people to share the cultural and natural significance of their traditional country.

Chapter summary

A systematic search of the literature identified eight case examples in which primary stakeholders were involved in the design and use of outcome indicators and evaluation use. The examples demonstrated a range of different approaches all of which fall under the broad term of participatory evaluation.

Although it had been hoped that existing criteria such as that provided by Guijt (2000) and Reed et al. (2006) could be used to classify the cases examples, the analysis of the cases revealed a broader set of conditions under which primary stakeholders were able to be involved. Therefore, in this analysis, four phases in which primary and other stakeholders could be involved in the design and use of outcome indicators, and evaluation use were identified: planning, indicator design and use, data handling, and evaluation use. It would appear, based on the findings in the reviewed cases, that often primary stakeholder involvement and participation was in phase two activities - the design of outcome indicators, but few were involved in the phases of planning, data handling and evaluation use.

A variety of characteristics of indicators emerged that made it possible for a framework to be developed for understanding the complexity of the interrelationships that exist between indicators and goals and objectives, negotiations amongst and within stakeholder groups, indicator use and quality. The section also presented the nature of indicators. They tend to be culturally appropriate, simple, visual, qualitative and/or quantitative, and location specific. The review found that there appears to be no formula in terms of the types of indicators designed. What is apparent is that each case uses the material and knowledge available to them during such processes.

There was discussion on the impact of facilitation, culture, participation, power dynamics, time and cost in relation to involvement of primary stakeholders in the design and use of indicators and evaluations. The importance of having a skilled facilitator in such processes points to the importance of having skilled facilitators in the design of indicators. Having a team that does not totally comprehend the differences between, for example goals, objectives, indicators, and measures can lead to disillusionment by all those involved and affect future processes (Hillyer & Purohit, 2007).

Achievement of empowerment and appropriate plans and indicators were seen as important outcomes in the case examples. The term empowerment is at times used loosely without adequate explanation of how this was determined. A table based on Vernooy et al.'s (2006) definition of empowerment in participatory monitoring and evaluation is developed and used cautiously to classify the level of empowerment in each of the eight case examples.

Chapter Five: Local case study

Introduction

This chapter presents the local case study. In the first instance, Part One sets the scene by outlining briefly the New Zealand context within which the fieldwork was carried out. Part Two presents the findings from the local case study research.

Part One: Context

Background

Different approaches are used for evaluating community development interventions in New Zealand. Although there are cases available online where evaluations have been conducted (Community Waitakere, 2012a, p. 89; Department of Internal Affairs, 2011; Evans & Clark, 2012; Schischka, 2009; Warsame, Mortensen, & Janif, 2014), there are relatively much fewer cases in which participatory evaluation approaches have been used (Schischka, 2009; Warsame et al., 2014). No cases could be found of evaluation or research processes in which primary stakeholders have been involved in the design of outcome indicators for development interventions. This section outlines the New Zealand context within which the local case study research took place. It starts by giving examples of cases in which community members have been involved in designing sustainability, cultural and community indicators that are not for a particular development intervention. The section then moves on to evaluation and indicator design in development interventions. Some examples of locally available guidance on indicator design are presented before a synopsis of Results-Based Accountability, widely used for indicator design for development interventions in New Zealand, is outlined. Part One concludes with a brief background on the organisation used as the local case study.

Participatory indicator design

Participatory indicator design in New Zealand appears to occur mostly when sustainability, cultural and community indicators are being designed. These indicators tend to

be developed in order to track changes due to a variety of influences in a particular locality over a period of time. The examples of participatory indicator design include involvement of community members in local government plans, Māori cultural indicators, and the development of community indicators.

Local government plans

Until 2010 local governments throughout New Zealand were required to produce a Long Term Council Community Plan that contained strategic direction for each authority for at least 10 years (Government of New Zealand, 2002). The Labour Government that introduced the Local Government Act 2002 sought to, amongst other goals, increase accountability of local authorities to their communities and "for local authorities to play a broad role in meeting the current and future needs of their communities" (Government of New Zealand, 2002, p. 23). As such community consultation took place every six years to identify the cultural, environmental and socio-economic outcomes that communities aspired to. Another obligation contained in the LGA 2002 was for local authorities to report three-yearly on progress made on achieving community outcomes based on measures and indicators developed. Leonard and Memon (2008) state that community participation in the process of agreeing on goals, outcomes, indicators, and reporting was largely restricted to consultation and written submissions, and that involvement of community members in the full process occurred in isolated cases. In a study conducted by Memon and Johnston (2008, p. 86) it was found that "relatively few local authorities in New Zealand have had significant levels of community involvement in developing their community indicators and monitoring and reporting regimes". They attribute this lack of involvement to "cognitive institutional barriers" (Memon & Johnston, 2008, p. 86) where design of community indicators was not seen as part of the process of developing outcomes. The requirement to involve community in the process of devising Long Term Council Community Plans ended when the Local Government Act 2002 was amended by the National Government in 2010 such that community participation was no longer required in the process of designing what is now simply a Long Term Plan.

Māori cultural indicators

Cultural indicators not only provide data on the state of the environment from Māori cultural perspectives, but also express Māori values, provide information on changes of cultural perspectives over time, and communicate their role in environmental monitoring

(Harmsworth, Young, Walker, Clapcott, & James, 2011). There is a body of literature that speaks to the involvement of Māori whanau (families), hapu (clans) and iwi (tribes) in the design of cultural indicators, especially for the environment. Nelson and Tipa (2012, p. 3) provide a concise overview of cultural indicators for monitoring "cultural health and wellbeing of significant sites, natural resources and environments".

In Auckland, the Independent Māori Statutory Board worked with various stakeholders, including iwi and hapu, to create indicators for cultural, social, economic and environmental wellbeing that are part of The Māori Plan (Independent Māori Statutory Board, n.d). The Māori Plan, developed between 2011 and 2012, contains five key elements that are Māori values, key directions, domains and focus areas, Māori outcomes, and indicators (Independent Māori Statutory Board, n.d). The 111 "'state' of wellness indicators' contained in the plan are meant to enable monitoring of progress, or lack thereof, in working towards development of vibrant communities, enhanced leadership and participation, improved quality of life, promotion of Māori identity, and a sustainable future (Independent Māori Statutory Board, n.d, p. 23).

Make it Happen Te Hiku

An example of community members and organisations working together to design community indicators is the Make It Happen Te Hiku project. The project started in 2013 to engage the wider community of Te Hiku in prioritisation of improvements in social services for the area. Its Project Action Group comprises representatives from social service providers, worked with multiple stakeholders, including local organisations, iwi, hapu and central and local government, to formulate community outcomes, strategies and actions for social development. A report was produced to communicate the priority outcomes for the community, strategies and actions that can be implemented by stakeholders to contribute to the achievement of these outcomes, and to propose a collaborative approach to improving community outcomes.

Five outcomes and ten indicators were identified in a consultation process during which 1,250 community members took part (Make it Happen Te Hiku, 2014). Figure 5.1 shows the community outcomes and indicators.

Figure 5.1: Make it Happen Te Hiku community and outcome indicators

Community Outcomes and Indicators (Summary)

This table shows the 10 indicators we will use to measure our progress.
The arrows show the direction in which we want change to happen.

Culturally strong, rich and proud	Peaceful, safe and connected	Prosperous and progressive	Clean and has a sustainable environment	Healthy and well
❶ Increase rate of enrolment in Kura Kaupapa ↑	❸ Decrease rate of Substantiated Child Abuse investigations ↓	❺ Increase Median Household incomes ↑	❼ Improve River Water Quality ↑	❾ Decrease incidence of Rheumatic Fever ↓
❷ Increase rate of achievement for NCEA Level 2: Te Reo Māori ↑	❹ Decrease rate of Youth Offending ↓	❻ Increase rate of Māori enrolment in Early Childhood Education ↑	❽ Increase Kiwi Wellbeing (Call counts) ↑	❿ Decrease the percentage of Cigarette (Tobacco) Smokers ↓

Response to the Minister for Social Development | April 2014

Source: Make it Happen Te Hiku, 2014.

RBA, discussed further in this chapter, will be the framework within which these indicators are used.

Indicator design for development interventions

The search for case examples in which primary stakeholders were involved in the design and use of outcome indicators, and evaluation use for a particular development intervention did not produce any from New Zealand. What was found, however, reflects attempts by local organisations to place value on evaluation and increase capacity of organisations in the design of indicators for use in determining outcomes realised through interventions. Examples of these are the work done by Community Waitakere in evaluation and design of indicators, and Adult and Community Education (ACE) Aotearoa in working with Māori and Pasifika to determine success as they perceive it.

Community Waitakere

Community Waitakere, a community development organisation in West Auckland, undertook an extensive research project that included identifying: the strengths, aspirations, needs and opportunities of the Waitakere community; evaluation practices; and indicators for community projects in 2012 and 2013. One of the key themes that emerged was that evaluation is not strength of the organisations working in Waitakere, but rather a contentious space in which it is unclear how organisations are held accountable and what indicators are used when measuring development (Community Waitakere, 2013a). In 2012, Community Waitakere conducted a retrospective impact analysis on projects delivered from about 2003 to 2013 in response to calls for greater accountability to multiple stakeholders, and more robust evaluation and documentation processes that showed the links from inputs to "outputs, outcomes, and demonstratable [*sic*] results" (Community Waitakere, 2013b). There is no documentation that suggests that the evaluations were participatory and inclusive of primary stakeholders. Nor is there indication that five case studies whose evaluation practices were documented (Community Waitakere, 2012a) involved primary stakeholders in participatory processes.

Community Waitakere's research project also collated a suite of social and community wellbeing indicators for use by community development organisations (Community Waitakere, 2012b). Of the nine domains in the suite of indicators, none were designed by local communities or primary stakeholders in local community development interventions. Rather they were chosen or adapted from existing indicators in New Zealand, the United States of America and Europe (Community Waitakere, 2012b).

Adult and Community Education (ACE) Aotearoa

Although not focussed on a particular intervention, and not aimed at designing outcome indicators, it is worth mentioning two separate research projects commissioned by ACE Aotearoa, an organisation involved in adult and community education. ACE Aotearoa commissioned research into the definition of success as defined by Pasifika and Māori in two separate projects (Adult and Community Education Aotearoa, 2014a, 2014b). Although the focus was on determining success with regard to "articulating a collective conceptualisation of 'literacy' and 'success' and the link between the two from the perspective of Pasifika peoples in Aotearoa" (Adult and Community Education Aotearoa, 2014b, p. 7), the aspirations and insights presented in the findings are not limited to the field of education. Rather, they are

presented in a manner that makes it possible for organisations working with Māori and Pasifika to turn the insights reported on into outcome indicators for development and other interventions. Guidelines for such an exercise are available locally.

Guidance on creating indicators

Duignan et al. (2003), Trotman (2008), and the Advisory Committee on Official Statistics (2009) provide guidance on design and use of indicators for community projects. The *Community Project Indicator Framework* (Duignan et al., 2003) provides a framework for funders and community organisations to use in determining the impact of projects. No assistance is provided in any of the documents on how primary stakeholders can or should be involved in the process of designing and/or selecting indicators. Instead Duignan et al. (2003) focus on the relationship and negotiations between a funder and a provider in choosing indicators. The focus of Statistics New Zealand's Advisory Committee on Official Statistics (2009) guidelines appears to be on sustainability indicators. Trotman (2008) provides direction on the whole evaluation process including on participatory processes. Part of the advice offered by Trotman is on the use of Results-Based Accountability within evaluation.

Results-Based Accountability (RBA)

Since the 1960s central and local government in New Zealand have been involved in supporting community development by providing funding and advisory services (Loomis, 2011). Over the past few years the RBA approach has gained popularity through the influence of government departments and territorial authorities.

RBA is "a disciplined way of thinking and taking action that can be used to improve the quality of life in communities, cities, ... and nations ... [and] can also be used to improve the performance of programs, agencies and service systems (Friedman, 2005, p. 11). It has been described as "a simple, common sense framework which communities, agencies and teams can use to focus on results (or outcomes) to make a positive change for communities, whanau and clients" (Ministry of Social Development, n.d.)**.** According to Friedman (2005), the author of RBA, the framework has two components: population accountability and performance accountability. At population accountability level, indicators are designed to measure how well agencies and organisations have done in terms of enhancing wellbeing at

community, city or national level, while performance accountability reports on the results or outcomes for clients of an individual programme or organisation (Friedman, 2005).

For performance accountability[1], RBA seeks to answer the questions: "How much did we do? [quantity]; How well did we do it? [quality] and; Is anyone better off?[2] [results]" (Friedman, 2005, p. 67). Friedman (2005) lays out a seven step process that begins by identifying the ends and works backwards to the means as depicted in Figure 5.2.

Figure 5.2: The 7 performance accountability questions

The 7 Performance Accountability Questions

1. Who are our customers?

2. How can we measure if our customers are better off?

3. How can we measure if we are delivering services well ?

4. How are we doing on the most important of these measures?

5. Who are the partners that have a role to play in doing better?

6. What works to do better, including no-cost and low-cost ideas?

7. What do we propose to do?
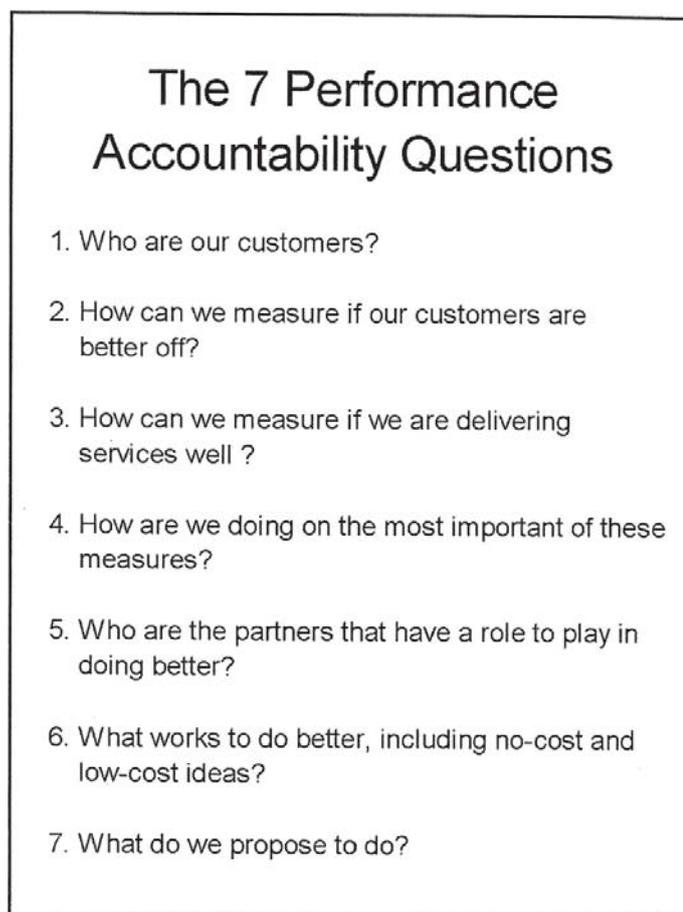
Source: Friedman, 2005 p83.

According to Friedman (2005, p. 130) "managers and community partners can use Results Accountability methods to do for themselves what evaluation companies are paid a lot

---

[1] The focus of this thesis is on indicator design at intervention level, therefore discussion on Results-Based Accountability is focused on performance accountability.
[2] Another question asked instead of 'Is anyone better off?' is 'What difference did it make?' This is the question that will be used in this study because the local case study prefers to use that question.

of money to do, including gathering data, analysing data and using data to improve performance." He discourages external evaluations stating that they are expensive, do not necessarily build the capacity of the organisation to conduct evaluations, can collect more data than the organisation needs, and at times are seen to be more important than the organisation or programme being evaluated (Friedman, 2005). Friedman instead encourages use of empowerment evaluation described as "the use of evaluation concepts, techniques, and findings to foster improvement and self-determination … [in which] program participants, including clients, conduct their own evaluations … to understand what is going on … from the participant's own perspective" (Fetterman, 2002, p. 89), that he claims is in keeping with the aspirations of RBA.

RBA is not without criticism. The claim made by Friedman (2005) that looking at an RBA report is like looking in a mirror because people see whether the community's quality of life is getting better or worse, is challenged. Keevers, Treleaven, Sykes, and Darcy (2012) contend that 'facts' are given privilege over 'value' "by representationalist views of knowledge … threaded through RBA" (Keevers et al., 2012, p. 11). The claim that RBA provides a simple framework for accountability is also contested, especially for small organisations that have to invest time and resources; detracting from the work they should be doing (Handley, 2008).

Keevers et al. (2012, p. 12) propose that, by giving precedence to measurement of service provision, "RBA excludes from mattering all the practices that cannot be constituted as a 'service'" such as 'giving back', working with people to be recognised and to have a sense of belonging. Data captured in the framework is based on quantitative measures for individuals and entities, and does not give credence to, for instance, relationships between people (Keevers et al., 2012). This leads to organisational practices in community development work not being visible to funders.

RBA continues to be the preferred model of accountability for contracts and grants given out by some central government departments and territorial authorities, examples of which include the Christchurch City Council (2010) and the Ministry of Social Development (n.d.). Organisations receiving financial support through government agencies are generally expected to report using this framework.

Organisation X

Organisation X, the organisation of the local case study, works with women, children, and men to provide family violence prevention, intervention and crisis management programmes and services in New Zealand. Supporting over 200 families each year, the organisation provides targeted interventions to address the needs of their clients. Organisation X has been operating in New Zealand for over 25 years. It's partners in service include families, communities, mana whenua (local Māori authority), Family Violence Services, Criminal Justice Services, Child, Youth and Family Services, health services, employers, cultural services, leisure services and education and training services (Organisation X, 2012). The goals of Organisation X are: the promotion of social justice and public understanding of, and concern and responsibility for family violence; provision of specialist services and development of community capability for families and communities to overcome family violence; and the development of strategic, organisational and financial capability within the organisation in order to achieve these goals. Amongst its objectives, Organisation X aspires to:

- create opportunities for people who have overcome family violence to encourage, inspire and support others to do the same;
- involve people with lived experience of family violence in service development and delivery; and
- develop measures, tools and technical capabilities for monitoring and evaluating of service effectiveness and outcomes.

Past evaluations

One of organisation X's aims is to create a culture of reflective practice (KI:1 meeting, 5 September 2013). The organisation has, in the recent past, commissioned two evaluations.[3] Neither of the evaluations involved clients in a participatory manner. Apart from the two evaluations the organisation does conduct evaluations of its programmes and operations as a matter of course KI:1, diary entry, 14 October 2013).

---

[3] In order to protect Organisation X's anonymity, the two programmes for which evaluations were conducted are not written about in this thesis as they are, to a large extent, unique in the family violence sector.

Introducing RBA

In July 2013 Organisation X started a process to introduce RBA into the organisation. The reasons for doing this were two-fold. On the one hand, Ministry of Social Development had for years been talking about use of this particular tool for evaluation, accountability and reflection; and on the other, the organisation was keen to ensure that it introduced RBA with measures that it had self-selected. They hoped to work with their clients on developing such measures in order to further the organisation's value proposition that the culture of the organisation becomes one that hears the voice of its clients and is led by it (KI:2, meeting, Diary entry, 14 October, 2013).

A consultant, working with Organisation X at the time volunteered to assist in the process because of her experience in two previous work places of developing quality frameworks[4], and her understanding of RBA. The consultant started working with staff through a two-hour workshop to introduce RBA by explaining what it is, what the organisation is trying to achieve, and why. After this workshop, the consultant met with senior staff to come up with measures for the first two RBA questions "how much did we do?" and "how well did we do it?" It was while trying to do this that staff became increasingly aware that a conversation needed to be had with clients, especially in answering the third question relating to "what difference did it make?" (KI:1, Diary entry, 05 September, 2013). The idea to involve clients by working with myself, a researcher keen on observing participatory indicator design was mooted to the organisation as explained in Chapter Four.

Part Two: Findings

Introduction

Part Two presents findings from the fieldwork. The case study used, as explained in Chapter Three, was selected for methodological considerations – it made it possible to look at the design of indicators in a 'real-life context' even though the intervention is not a

---

[4] A quality management framework is "the strategy, advice, guidance and tools necessary for an organization ... to attain quality, efficiency, and effectiveness in performing its mission responsibilities" (World Meteorological Organization, 2011).

development intervention. The analytical focus therefore is on Organisation X's process of coming up with RBA measures with its clients rather than the services and programmes it provides. The findings from the case examples reviewed in Chapter Four have framed the way in which the local case study data was analysed. The headings under which data is presented in Chapter Four are followed as closely as possible in Chapter Five. However, this is not appropriate in all cases as there is no data from the local case study for some of the headings used in Chapter Four. The term 'measures' is used instead of 'indicators' for two reasons. The first is that most staff use the term 'measures' for RBA, and RBA uses the terms 'measures' and 'indicators' to mean two different things. It is therefore appropriate to use the term measures in relation to what Organisation X was designing (except where staff are quoted using the term indicators).

The first section of Part Two looks at the phases of involvement of stakeholders and identifies the different roles that stakeholders were involved in. The next section looks at the characteristics of the 'measures' designed by Organisation X's primary stakeholders, their clients. The third section examines the dynamics of involving clients in the design of measures, followed by a section on what the clients and staff felt was achieved. The last section presents the aspirations of clients and staff in relation to the next steps to be taken in the process. A summary concludes the chapter.

Phases of involvement in participatory evaluation practice

The four phases of participatory evaluation practice identified in the literature and cse examples are planning, indicator design and use, data handling, and evaluation use. The discussion on involvement of clients of Organisation X in the process of coming up with RBA measures is based on these. This section does not have an 'Evaluation use' heading as there is no data available.

Phase 1 - Planning

The planning phase of the process to develop RBA measures involved staff, a volunteer consultant and me, as a participant researcher.

A steering group was set up in November 2013 and met three times. Its membership comprised KI:2, the consultant, a senior manager, an ordinary member of staff, and me. While KI:2, the consultant and I attended all meetings, other staff attended on invitation or when

there was no pressure of work. Although it had initially been thought that there would be clients on the steering group this did not eventuate. This is further discussed in the facilitation section.

The steering group made decisions about: the involvement of clients in the design of outcome measures for RBA; the decision to have focus groups to design measures; the nature of the measures to be collected through the focus groups; and the possibility of fitting measures into Te Whare Tapa Wha and Five Ways to Wellbeing, further discussed in the data handing section.

Phase 2 – Measure design and use

Four focus groups were held to involve clients in the design of outcome measures for RBA. Each focus group was held to cater in different ways to clients that were participating in programmes. One focus group was for clients that were training to be peer support specialists, two for clients who were attending educational classes on family violence, and one for Māori clients.

At each of the focus groups the facilitators explained that the measures to be designed for Organisation X's future evaluations, or analysis as one staff member called it, would begin with a conversation about the clients' feelings. These feelings would be based on how clients had felt during the period they experienced family violence; how they were feeling when accessing help from Organisation X or any other organisation that addresses family violence; and then how they are feeling or aspire to feel when they are leading violence-free lives. These feelings came to be known collectively as the Feelings[5] and are listed in Figure 5.4 (see the Characteristics of measures section).

Phase 3 – Data handling

The steering group met after all the data had been collected to decide on next steps in the process. It was suggested, and transpired, that the Feelings be put together in groups to identify those that were similar. They would then be fitted into wellbeing models. In discussing the rationale for fitting the Feelings into wellbeing frameworks, one staff member noted:

---

[5] The words and phrases listed by clients when discussing their feelings are referred to as 'the Feelings' in this thesis.

Both are wellbeing models rather than violence models … we've heard what women have said in terms of their aspirations and dreams. We can't make violence our measure. We have to make wellbeing our measure (Interview: S3).

A staff and steering group workshop was held. One staff member presented Te Whare Tapa Wha and Five Ways to Wellbeing[6], two wellbeing models, so that the group could consider whether the Feelings could fit within the frameworks.

A suggestion made (WP2) was the possibility of achieving Five Ways to Wellbeing through Te Whare Tapa Wha. This resulted in some discussion and effort to align the two as depicted in Table 5.1.

Table 5.1: Alignment of Five Ways to Wellbeing with Te Whare Tapa Wha

Te Whare Tapa Wha	Five Ways to Wellbeing
Taha tinana (physical)	Be Active
Taha wairua (spiritual)	Take notice Give
Taha whanau (social)	Connect
Taha hinekaro (emotional)	Keep Learning

Source: Author.

Discussion on this aspect with research participants brought out the following sentiments:

I struggled initially with this but when we did it [staff workshop] as staff it began … to make a little more sense … I would have liked to … see whether there was that translation from feelings to framework and how appropriate the frameworks were (Interview: S1).

I think it's trying to fit something into a box … you will find that people from one [cultural] group are not even going to be able to relate to what's in one model (Interview: S4).

---

[6] Te Whare Tapa Wha is a holistic Māori wellbeing model that recognises the four cornerstones of health as *taha wairua* (spiritual), the pertinence of identity in relation to culture and connections with others, ancestors, and the natural environment; *taha hinekaro* (emotion), Māori ways of feeling, thinking and conducting oneself; *taha tinana* (physical), bodily health; and *taha whanau* (Ministry of Social Development), social wellbeing (Durie, 2011). Five Ways to Wellbeing are: connect with other people; be active through physical activity; take notice of what is going on around you; keep learning about anything that captures your interest; and give whatever you have including time, smiles, kind words to those around you (New Economics Foundation, 2008). The two models are further described in Appendix Five.

The content of the staff workshop was considered a trial run to see what the result would be, before clients would be asked to go through the same process. At the end of the workshop, the general consensus was that it is possible to fit the Feelings into the wellbeing models, but it would be up to the clients to decide how the Feelings would be grouped into one or both the models.

All the clients interviewed said that putting the Feelings into wellbeing models would give them structure. It was found that all four had no or limited knowledge of what Five Ways to Wellbeing is, while two Maori women knew and understood Te Whare Tapa Wha. C1 and C4 (Interviews) suggested that Te Whare Tapa Wha could be used because from a Māori perspective, the four cornerstones are what people who have experienced family violence look forward to achieving. None had knowledge of how RBA works, as this had not been explained.

Turning Feelings into measures

Although Organisation X had already developed draft RBA measures for the first two questions, these were not used to inform discussions at steering group meetings nor focus groups. An example of the measures is Figure 5.1.

Figure 5.3: Organisation X RBA scorecard

RBA Scorecard: Individual and group programmes: Women and children

Client Groups: Women aged 18+ who have experienced family violence and children up to age 12 who have experienced (including witnessed) family violence.

Service Intention: To increase women's and children's client's knowledge and understanding about family violence and facilitate personal choices and behaviours that sustain personal safety and wellbeing.

Outcome: Women and children are living fulfilled lives free from violence.

Performance Measures

1. **How much do we do?**
   - # of women and children who begin the programme
   - % of women and children on programme with Protection Orders

2. **How well do we do it?**
   - % of women and children who fully complete the programme
   - Client-reported experiences of the programme

3. **Is anyone better off?**
   Positive improvements in safety (indicated by repeat incidents of violence) and changes in knowledge, beliefs and skills leading to quality of life improvements. Time intervals: upon file closure, 3 & 6 months.

Source: Organisation X Draft RBA measures, 2014.

The next step is to turn the Feelings into measures that can be used for RBA. Part of the process of turning Feelings into measures is to include a workshop with clients to talk about firstly putting the Feelings into groups to be fitted into the two wellbeing models, and secondly deciding how the two models will fit into RBA. Because the process of turning the Feelings into measures had not been completed, staff and clients were asked how they think this will be achieved. S3 was clear on how this would work suggesting:

> We know that people who are socially isolated are more vulnerable to violence, and those who use violence are more likely to ... Through the Five Ways to Wellbeing social connection ... is important .... so developing an RBA measure that identifies whether or not people are feeling more connected will provide us with important information about the extent to which they remain vulnerable to violence (Interview: S3).

Other staff members interviewed said thought turning the Feelings into RBA measures would be challenging because:

> We were looking at tools [wellbeing models] to help us process that data in a way that gives us indicators ... What is the link between the two [wellbeing models and Feelings] so that it's an outcome measure? ... how, from the ground up, do we make the links in a realistic, not airy fairy kind of 'this is what we are aiming for' but seriously, practically, literally, how do we achieve that? (Interview: S1).

> That is going to be the tension between doing this process and trying to put it into ... MSD's [Ministry of Social Development's] RBA outcomes ... But at some point ... [we have] to be able to say 'can we measure things like how they [clients] got a better understanding of what family violence is and how it impacts on them (Interview: S4).

Similar views were expressed by clients about turning Feelings into RBA measures, with one saying:

> I guess it's really hard for me. I've come from a maths analyst side of things ... So measuring words is a really tricky thing to do because we all love stats don't

we? But then stats don't mean a lot ... And I guess the feelings are what are important because you are dealing with a life (Interview: C2).

Characteristics of measures

The phase during which measures were being designed involved both staff and clients. It was the clients, though, who came up with the Feelings that would be used in determining measures for the last RBA question, "What difference did it make?, and that is the focus of this section. It looks at how the Feelings were designed, their links to the organisation's goals, the negotiations that took place to arrive at them, the nature of the Feelings, and their quality.

Figure 5.4 brings together the Feelings from the four focus groups. The presented merges the Feelings captured by the facilitators and those that I, as an observer, noted from the discussions.

Figure 5.4: The Feelings

Before intervention	When accessing services	Future aspirations
Alienated	Able to be by myself	Attractive
Alone	Accepted	Be more myself
Angry	Accomplishment	Beautiful
Can't cope	Achievement	Brave
Desperate	Alienated	Bubbly
Failure	Alone	Connected
Fear	Being given a chance	Can make choices
Frustrated	Belief	Determination
Guilt	Bubbly	Educated
Helpless	Can prioritise	Empowered
Hopeless	Content	Excited
Insecure	Comfortable	Feminine
Isolated	Educated	Find myself
Lonely	Emotional	Free
Loss	Excited	Full of laughter
Lost	Growth	Happy
Low esteem	Happy	Have life tools
No self esteem	Have goals and dreams	Have normal relationships
Nervous	Have choices	High self-esteem
No expectations	Helpless	Independent
Overwhelmed	I'm working towards something	In control of my life
Paranoid	It's real	Inner strength
Sad	Learning	I own my body
Scared	Less fearful	Like I have a voice
Shame	Listened to	Listen to
Stressed out	Motivated	Mana (honour, authority)
Tired	Not financially stressed	Not living on egg shells
Unsafe	Not judged	Optimistic
Unsure	People interested in me	Peaceful
Very uncomfortable	Provide safe & happy environment for children	Positive
Vulnerable	Realising I am not the only one	Proud
Weary	Relieved	Safe
Without choices	Safe	Self-caring

Worried Worthless	Safe in the group Strength Stressed out Supported Understood Worrying Weak Welcomed	Self-confident Sexy Shame-free Understood Waiora (wellbeing) Wearing what I like

Source: Organisation X and Author.

Linked to goals

The goals of the organisation have been presented in Part One. Although the organisation goals were not discussed by clients and staff during meetings and focus groups, interviewees suggested that there was alignment between the opinions the clients voiced in relation to their future aspirations and the stated goals of Organisation X.

I see the links between our Feelings and the goals … there was a lot of loneliness, hurt … sadness … the next step was my goals and dreams. Organisation X has the same goals and dreams (Interview: C4).

The sentiments expressed by staff were similar. They talked about getting the alignment between organisational and individual goals right.

It's hard to set out your vision and objectives and know that that's going to fit 100% with what clients really, really want. We didn't start from a basis of being consumer-consultative when we have looked at outcomes and RBAs. But looking at what the women wanted … they're the outcomes we're trying to achieve (Interview: S2).

Negotiated

During the focus groups, in order to come up with the Feelings, the clients expressed what they personally wanted to communicate and these became 'statements on the table' that were often then generally discussed by the group. There were no opposing views, or negotiations. Women generally agreed with the words other women voiced and often affirmed these with their own stories. Staff members were not involved in these discussions other than as facilitators. An example of discussions follows:

Angry! Very angry! (FG1:2).

Same. He broke my hand and then blamed me for it. I wanted to beat the living shit out of him. Big bully. Anger came because of the psychological stuff he put me through. I wasn't allowed to express my feelings. I wanted to put a knife in his back while he slept one night. I was so angry (FG1:4).

My anger was because I kept asking myself 'why am I here?' I was mad more at myself than him (FG1:3).

There were, however, numerous times when women discussed some of these feelings and worked together to identify words that would articulate the feelings they wanted to convey. An example of this is:

Insecure. Vulnerable. He had me so controlled because I had no bank account … Every time I came here I got more confirmation of how abnormal this stuff was. I had normalised it. Vulnerable (FG2:4).

I too felt vulnerable and insecure. So yes let's add that to our list. But when you spoke I thought what you talked about was control as well … Do you think? … He took over my life and controlled me (FG2:1).

Yes it is about control isn't it? (FG2:4)

Two focus groups had two participants. The clients mostly spoke in turn rather than discussed their feelings.

Nature of measures

Staff and clients commented mostly on the Feelings in relation to them being culturally-based, not simple, qualitative and quantitative, and location and context specific.

Culturally appropriate

To ensure that the measures designed for use in RBA were culturally appropriate for most clients, a focus group for Māori women was held. This recognised that colonised women bear the brunt of family violence more than others (S1: steering group meeting, Diary entry, 18 November, 2013). Asked why she took part in the focus group for Māori women, C4 explained:

I knew that this was something from a Māori woman's perspective and that was the driving force for me (Interview: C4).

The feelings that the Māori women talked about were similar to those that came out of the other three focus groups. In order to gain a better understanding of how the organisation could better work with Māori women, the women were also asked about the practical support they needed, what they hoped to achieve in life, and what they would say to encourage other women to access services.

(Not) Simple

When talking about the Feelings and whether or not staff and clients thought they were simple, most did not understand how the Feelings would become RBA measures and felt that this made the Feelings and the process complex (Interviews: S1, S2, C2 and C4). One client stated:

I think it could have been better explained because I … I do wonder in the back of my mind how different my feedback could have been … I don't think I would have been focussed on the Feelings. I probably would have talked more around [measures for] supports in that before, during and after stuff instead of past, present and future Feelings (Interview: C4).

Qualitative and quantitative

Because Organisation X would like to measure how much they do, how well they do it, and what difference it makes, both qualitative and quantitative measures are to be used (Interview: S3 and S4). The feelings of women are to provide qualitative data. One client commented:

These words, it's great … how the hell do we measure this? Coz you can't put where you want to be in numbers in this situation (Interview: C2).

Location and context specific

Staff members of Organisation X said that although the measures being designed by the women are currently specific to their organisation (Interviews: S1, S2 and S4), there is the ability to be responsive to best practice (Interview: S2), and to be able to do comparative work:

> You need to be able, to ... do comparative work ... I think as always there is a place for collaboration ... people don't have a conversation about 'how do we develop these things together with communities?' ... there will be common denominators (Interview: S4).

> I think that you have to go for those generic wellbeing frameworks like Te Whare Tapa Wha or Five Ways to Wellbeing, because they can then be applied to different sites ... we're then in a position to know across and within organisations that provide multiple services, what [services] are most effectively contributing to achieving that (Interview: S3).

Quality of indicators

By the time fieldwork came to an end, the measures had not been tested or used in any way. Most staff members were not clear on how the Feeling related to the models as expressed in the workshop:

> How are the Feelings, placed in that [wellbeing models] going to measure safety, and increased knowledge about violence using RBA? I don't get it because I think that 'being active' for instance is work for other organisations to do, not ours. I think we need measures that we can use for our programmes (WP5).

> There are indicators that need to be defined by the agency. For example safety is one of our key issues. So we need to highlight this. The measures women have come up with don't cover everything. Will we have a conversation with them or on our own to include such measures? (WP1).

Dynamics of involvement and lessons learned

Clients and staff talked positively about their experiences of the participatory development of measures. They also spoke about aspects that could have been improved on. This section offers an account of the dynamics of involving clients and the lessons learned, including facilitation; participation; power dynamics within and between stakeholder groups; cultural considerations; time; and cost.

Facilitation

The four focus groups were facilitated by three staff and one client[7]. Facilitators felt that they had achieved what they set out to by engaging women and being able to draw out the data sought (Interviews: S1 and C1). A staff member commented:

The consistency across the groups … I don't know whether it means you're asking the right questions or your questions are limiting to make sure you get the answer you want … I think that is certainly one thing we did do well (Interview: S1).

There appeared to be no prior preparation by and for facilitators for each focus group. An example of this was a conversation with two facilitators before a focus group. During the drive they asked me to talk to them about how the other focus groups had been facilitated and the questions asked, and some tips about conducting focus groups. I cautioned them against asking questions about programme delivery as this had not been discussed in the other focus groups.

From a client perspective, there was the view that the facilitators had done a great job of drawing out discussion and making women feel comfortable enough to share what they wanted to (Interviews: C1, C2, and C3). This aligned with the viewed expressed by S5 (Interview) that mutual comfort appeared to have established between facilitators and clients. They found the discussions to be important for reflecting on their journeys by identifying how they had felt during the three phases helped them to see how far they had come. Clients, however expressed that facilitators could have given more information about what the whole process would involve and what would happen to the data collected (Interviews: C2, C3 and C4).

Culture

A focus group was held for Māori women so that they could have a conversation that took into account anything that may have been important for them as Māori women (Interview: S3). During the staff workshop where discussions took place around Te Whare Tapa Wha and Five Ways to Wellbeing, it because apparent that there were differing perspectives based on cultural background.

---

[7] To increase anonymity of facilitators, they have not been identified in this section.

C1 and C4 (Interviews), both Māori women, expressed that it was important for them that the Feelings are contained in Te Whare Tapa Wha because they can relate to that from a cultural point of view. C1 (Interview) pointed out that she felt that cultural considerations should have been part of the process from the beginning.

C4 (Interview) said that by ensuring there was kai to share at the beginning, Organisation X had done the right thing in recognising a vital Māori cultural practice. She however noted there was not much time put into explaining the process but:

Culturally that's how we roll ... 'hui's on!' 'Ok, I'll be there (Interview: S4).

S1 and S3 (Interviews) stated their uncertainty in terms of whether the organisation could have done more to take cultural considerations into account. In her view, S1 felt the manner in which the focus groups had been facilitated was from a mainstream, Pakeha New Zealand perspective and possibly not enough thought was given to various aspects of culture that could have been taken into account.

Trying to fit Te Whare Tapa Wha into the Five Ways to Wellbeing was uncomfortable from a cultural perspective:

When we were trying to fit Feelings ... into Te Whare Tapa Wha ... at no point did anybody reverse it! ... people are more inclined to put what they know forward first ... I don't know if you remember me talking about why we were having so much hassle making it fit ... it's not actually meant ... And I voiced my concern that it was kind of feeling like the third version of the Treaty ... we have the Māori version, the pakeha version , and then we'll make up our own ... That's what was happening because if you think about it we were trying to make Te Whare Tapa Wha fit into 5 Ways to Wellbeing when it is actually 5 Ways to Wellbeing that should fit into Te Whare Tapa Wha (WP4).

Participation

This section presents findings on the power dynamics amongst and within stakeholder groups, and the manner of participation by clients. It does not, as in Chapter Four, address issues of enhancing quality of participation as this was not specifically mentioned in any discussions held by staff or the steering group.

Power dynamics

There were differing views on power dynamics in the process. While all staff said that power dynamics are always present, some clients did not feel this was the case during the focus groups.

Power dynamics amongst stakeholder groups

When asked about power dynamics amongst the different stakeholder groups i.e. staff, steering group and clients, there were different views offered by clients on the experiences they had in the focus groups. The clients had not (apart from two) taken part in the same focus groups. There were basically two opposing views. Sentiments expressed include:

> I did not feel uncomfortable with [senior staff] facilitating because she's the nicest person in the world … and she shared some of her own bits and pieces, so she's just one of us (Interview: C2).

> I do think there were power dynamics … When you're vulnerable, the girl in the [safe] house, and you don't quite have your wits about you, and then [senior staff] and these strangers come rocking in, you're not going to present how you might with a one-on-one type scenario … I just think when you've got [senior staff] part-time worker, person in safe house and another Māori woman that's coming in because there is a Māori focus group …and then lack of knowledge about actually why we're all here … it could have been [done] better (Interview: C4).

Staff said they were aware that there were power dynamics (Interviews: S1, S2, S3, and S4). Views expressed include:

> I think that there was quite quickly established mutual comfort and I think that was achieved perhaps through … trying to stand in other people's shoes and look at [senior staff] and trying in an appropriate way to manage that maybe through humour or … getting down on the floor (Interview: S3).

> I think you always have to be conscious of that when you're in a room full of people where you're asking them to talk about stuff that is very private to

them. Talking about times when they were very vulnerable, and talking about things that they may be feeling ashamed about (Interview: S2).

Another view expressed was the need to be mindful that clients also want to please staff, and that client participation can be due to power dynamic (Interview: S4). It was suggested that this could be mitigated by having coalface staff to facilitate the groups so that they hear directly about what clients want (Interview: S4).

Power dynamics within stakeholder groups

Clients all said that they felt comfortable with each other during the focus groups (Interviews: C1-C4). They said that between them they had not felt there was power at play because they felt as though they had treated each other as equals in the process.

Power dynamics were visible in the steering group. Because all staff acknowledged this to be the case, they said it was important to recognise that it is there and to be aware of the effects it may have on the process (Interviews: S1, S2, S3, S4). This was succinctly expressed:

> I think you get that where you have a lot of smart people coming together with some really strong ideas … different life experiences, different work experiences … competing views … that's very normal and healthy and I guess with those competing views you have a bit of power play. My view is more important, it's better informed, I have that experience, I have that knowledge (Interview: S2).

Another aspect that was spoken about is the competing views of operational staff and management, in which the agenda of the process was driven by senior staff:

> when we have some of these conversations you think you are headed in the same direction but you realise that on one level senior management has a different view to what you might be thinking on an operational day-to-day 'this is what we do' … Sometimes those strategic versus operational perspectives can clash (Interview: S1).

Manner of participation

Participation during the focus groups was described by everyone interviewed as lively. Words used to describe the manner of participation includes enthusiastic, non-judgmental, active, positive, honest, very real, fair, (Interviews: C1, C2, C3 and C4). There was a feeling that everyone wanted to be there and conversation flowed (Interview: C2) and that they could say anything (Interview: C3). I observed that there was a lot of emotion as clients spoke about their Feelings. Some laughed, there were serious moments, there were tears and there were determined looks. Clients expressed that facilitators ensured that they were mindful of the emotions expressed by clients and that time was taken to acknowledge the Feelings clients voiced and spoke about.

Staff views aligned with this assessment. There was a feeling that once clients got into discussion they felt comfortable enough to share honestly and openly to the extent that they possibly benefited from the discussion (Interviews: S1, S2 and S3).

Length of time

All those involved in the process and who were interviewed expressed the sentiment that the process had taken a long time.

Clients said that they understood that processes like this do take time especially if the organisation wants to get things right (C2, C3 and C4). There was alignment of views that there was a need for Organisation X to keep in touch with them so that they are aware of what is going on, stay interested, and know whether the process is ongoing or if they are still required to participate (Interviews: C1, C2, C3 and C4). A sentiment expressed was that the time since the focus groups meant that momentum had been lost and clients may have lost interest, moved on, or would not be able to remember what had been discussed in the focus groups (Interviews: C2 and C4).

> We had the focus group ages ago ... life does happen and I guess this far down its hard to hold on to that good juicy stuff that I could have given at the beginning of the year ... if they ask me a lot of that stuff now ... I don't think I'd feed it back the same (Interview: C4).

> I would say we've lost half [of the clients] because ... when we mail out to our peers ... you can guarantee you'll get [messages] back saying email has

changed. That to me is an indicator – gone! … [in past research] I'd receive a little email every now and then … just to say hi … but it still kept me informed to the point I felt included (Interview: C1).

Although there was between eight and 10 months since clients had been involved in the process, they all expressed the desire to continue to be involved (Interviews: C1, C2, C3 and C4). The reason given was that the organisation has done so much for the clients and they would like to contribute to a process that will help other women who experience family violence.

Staff views were comparable to those of clients. There was a feeling that momentum had been lost (Interviews: S1 and S4) although this was largely due to demands on time. There is recognition that the environment within which they work, and were trying to achieve design of measures for RBA, is a busy one. There was no prior recognition of the time that would be required for this process (Interview: S3) and that it was no surprise that it had taken the amount of time it had (Interviews: S1 and S2). Another aspect was this was a learning process for those involved and required more time than possibly future processes would take (Interview: S4). Concern was raised that often the time it takes to prepare is taken up when other demands on time are present, leading possibly to pitfalls of tokenism when engaging in the process (Interview: S2). S1 was also of the same opinion stating:

Perhaps it goes back to how much do we value this process, how much space are we willing to create to allow for this? (Interview: S1).

Suggestions put forward to make the process take less time was that the process would have possibly been shorter had there been a coordinator whose job it was to make the process happen (Interviews: S1 and S3) and to have a plan (Interview: S2).

I love when people can say to me '… this is the plan we're going to work towards' … you might be needed, what's going to be required of you … I think we really need to have a framework … know what we're doing, know what it's going to take … because the worst thing you can do is not plan and then do it wrong and have it be a negative experience for the person you've invited to participate, the staff or the organisation itself (Interview: S2).

Cost

The costs that went into the process on the part of Organisation X were providing food at the focus groups, stationery and staff time. Staff members were asked whether they thought that the cost that had gone into the process was commensurate with what was being achieved. The general view expressed was that if a cost-benefit analysis were to be done, the benefits would far outweigh the costs because the views of clients in such an important process matter (Interviews: S1, S2, and S3). S3 (Interview) stated that this was an investment and would only be considered a cost if there was no improvement made on what Organisation X wanted to achieve. There was recognition that staff had put time into the process instead of service delivery because this was important to the organisation (Interviews: S1, S2 and S4).

The next section attests to the importance of this process to the clients and staff because of what they feel was achieved.

Achievements

The achievements realised in the process of designing measures for RBA are described by staff and clients as being numerous. Clients and staff of Organisation X expressed that a lot had been achieved and that empowerment was one of the key triumphs. This section presents the clients' words to describe empowerment and reflections on their personal growth through the process. It also puts forward the measures and opportunity for clients to give back to the Organisation X's accomplishments.

Empowerment

As stated in Chapter Three, empowerment in participatory monitoring and evaluation occurs when people are put in charge, have their skills developed, note that their views count, and learn alongside others (Vernooy et al., 2006, p. 401). In this regard, clients felt that their views counted, and to some extent there was opportunity for joint learning.

Were primary stakeholders in charge?

The steering group was mostly in charge of the process of devising measures for RBA. No clients were on the steering group, nor were they involved in any decision making processes. Clients cannot therefore be said to have been in charge of the process.

Were the skills of primary stakeholders developed?

One client participated as a co-facilitator for two focus groups. She had the following to say about having her skills developed as a facilitator:

Any knowledge is good knowledge. Never having done that type of group, being called in last minute. Facing it as a challenge not a weakness. It wasn't until I was on the other side that I said 'wow, I just did that!' and in reflection at home I was so proud of the success … 'You go girl'" (Interview: C1).

Did the views of local people count?

The clients said they were happy to have been asked to take part in the design of measures to be used for evaluation within Organisation X.

That your opinion was valued. That you were being given an opportunity to improve the programmes going forward … that Organisation X was aware that they needed to keep looking at what they were doing to see, check, what was needed (Interview: C3).

Nice to feel as though you're giving a bit of input into things and hopefully then just fine tuning things for all the other women that are unfortunately going to have to use Organisation X for years and years to come (Interview: C2).

So I know that having been involved I have the capability of helping so many more and that's what drives me. It made me feel included and it made me as a survivor feel safe for other survivors, that their backs were covered (Interview: C1).

Staff expressed similar sentiments on the views of clients stating that they hoped clients felt heard (Interviews: S2 and S3) and that their views and participation was important to the organisation (Interview: S3).

Was there opportunity for joint learning?

Discussions on developing measures for RBA were held by staff and clients separately. The opportunity for joint learning was therefore not presented.

Reflection on personal growth

Apart from stating that it was important that their views counted, clients felt that coming together to reflect on their personal journeys had been another opportunity to learn about their own growth.

I guess it showed me my own personal growth because when I sat there with some of those people that I had been involved with at the front-end of my wellbeing and could identify [the Feelings], that prior to that I couldn't even identify … obviously a lot of stuff has happened for me, there has been a lot of movement, and my journey's been full (Interview: C4).

Sitting down there and having everything written down really made you think where you are, alert to where you want to be as well … And then that whole thing of 'this is what Organisation X can do' and actually it's done that, moved you from the negative things to a completely different life (Interview: C2).

Another sentiment expressed by C2 (telephone conversation, diary entry: 2 October 2014) was the motivation to enrol in a course at university because she could see how much progress she had made in her life, having reflected on this during and after the focus group.

Appropriate measures

Although the process of working with clients on turning Feelings into measures had not been completed, staff and clients felt that it would be appropriate to design indicators based on what clients aspire to feel in future (S3, C2 and C4), and also those that assist in the tracking of change based on service delivery (S1, S4, WP4 and WP6).

An opportunity to give back

Clients who participated in the focus groups did so with great enthusiasm. They did not have to be encouraged to speak. For all the clients (Interviews) this was their chance to give back to Organisation X, and to contribute to better services for those who are yet to unfortunately access services of Organisation X:

You wanted to give something back. I think that's where we were coming from that day because they had given us so much as well and by just talking to … if we could help somebody else get through the hump that we were in, then that

was great and I guess it was a good feeling for us as well to feel like you could do that (Interview: C2).

C4 (Interview) stated that she had attended the focus group not really aware of what the process was about but was motivated to give back to Organisation X and to be able to contribute to a conversation from a Māori woman's perspective.

Next Steps

All the clients interviewed stated that they would like to be part of the next steps in completing the process. A suggestion made by clients was for the Organisation X to get the focus groups back together so that clients could communicate their aspirations and to discuss with the organisation what the whole process of designing outcome measures for RBA would be, to enhance their understanding.

Aspirations of clients and staff for future involvement of clients

The participants pointed out (using Appendix Four), the aspects they felt were appropriate to be involved in when working with Organisation X in this process. Table 5.2 presents a snapshot of the answers given by the eight participants who took part in the semi-structured interviews, indicating whether or not clients should be involved in the phases and roles presented. In summary, most staff and clients felt that clients should, as far as possible participate in all phases, tasks and roles.

Table 5.2: Aspirations of clients and staff for future involvement of clients

Phase	Role	C1	C2	C3	C4	S1	S2	S3	S4
1 - Planning	Initiator	N	N	Y	Y	N	N	Y	N
	Research/evaluation team member	Y	N	Y	Y	Y	Y	Y	N
	Research/evaluation designer	Y	N	N	Y	Y	Y	Y	N
	Needs assessor	Y	N	Y	Y	Y	Y	Y	Y
	Trainer	Y	N	Y	Y	Y	Y	Y	N
2 - Indicator Design and Use	Data collectors	Y	Y	N	Y	Y	Y	Y	Y
	Indicator designers	Y	Y	Y	Y	Y	Y	Y	Y
	Indicators selectors	Y	Y	N	Y	Y	Y	Y	Y
	Indicator aggregators	Y	Y	N	Y	Y	Y	Y	Y
	Evaluators	Y	Y	Y	Y	Y	Y	Y	Y
3 - Data Handling	Data analysts	Y	Y	Y	Y	Y	Y	Y	N
	Report writers	Y	Y	N	Y	Y	Y	Y	N
	Information distributors	Y	Y	Y	Y	Y	Y	Y	Y
4 - Evaluation Use	Evaluation users	Y	Y	Y	Y	Y	Y	Y	Y
Key: Y = Clients should be involved N = Clients should not be involved.									

Source: Author.

Phase 1 – Planning

Five clients and staff felt that an organisation will probably be the initiator of the evaluation practice process, while three felt that if clients are involved routinely in the life of the organisation, they too could be initiators of the process. C2 and S4 both felt that most aspects of the planning phase should be left to the organisation, because it knows why it would like to embark on such a process. The other research participants felt that there would be value in staff and clients figuring this out together so that the data collected are relevant for both.

Phase 2 – Indicator design and use

Seven of the eight research participants said that clients should participate in all roles of the indicator design and use so that the indicators reflect the aspirations of clients, and that as evaluators, clients will be able to interview peers about their experiences. Only one client (C3) felt that clients should not be involved in all roles because their personal biases could affect outcomes.

Phase 3 – Data handling

In the data handling phase six participants said clients should be involved, because of the different interpretations that staff and clients could have on the data collected, and that clients would be able to work with others to decide what goes into the report. C3 and S4 felt that there are skills required for some aspects of data handling tasks that not all clients may possess. This phase should therefore be left to the organisation to do.

Phase 4 – Evaluation use

Lastly all participant views aligned on the need for clients to be involved in evaluation use.

Benefits and challenges of involving clients

Both clients and staff expressed the importance of involving clients in participatory design of measures for RBA including planning, indicator design and use, data handling, and evaluation use. The ability to leverage on the clients' lived experience of family violence, and other skills, knowledge and worldviews they can potentially bring to such processes was highlighted.

The benefits of involving clients seem to far outweigh the challenges according to staff and clients. C1 and S2 (Interviews) spoke about participation as a means to give clients the opportunity to be involved with something meaningful, feeling that they are in control and have something to give. C1, C2, C3 and C4 (Interviews) stated that the process could be enriched by involving those who have lived experience of family violence and are able to contribute from a different point of view to those who have academic understanding of it. S4 (Interviews) asserted that by involving clients in such processes, there is greater ownership of the process and increased likelihood in clients being engaged with services. Client participation also brings in alternative perspectives that are not necessarily aligned to what staff are used to (Interviews: S1 and S3).

A challenge that was talked about often was that of time. It was however felt that the method of coming up with evaluation processes can be prioritised, with sufficient time set aside for it (Interviews: C4, S1, S3 and S4). Another way in which time challenges could be overcome would be to appoint a coordinator of the process who will be responsible for coming up with a plan and time frames (Interviews: S1, S2 and S3). If not done properly, and with insufficient time being set aside, there is the trap of falling into tokenism (Interview: S2). It was also said that it is not normal to involve clients in such processes, and that those at the beginning of their journey to being free of family violence, may not be in the mental and emotional space to participate meaningfully (Interview: S4). Participatory measure development can also be a resource-intense process if done properly, because involving clients in a meaningful way requires the necessary supports to be put around them (Interviews: S2 and S4). Another challenge might be that the agency might feel loss of power and control that any agency may feel in any participatory evaluation process, and lose the ability to see that as a positive thing (Interview: S1).

Chapter summary

This chapter began by providing the context within which the local case study was conducted. It started by considering the effort put into participatory indicator design in the field of sustainability, cultural and community indicator design and noted the absence of similar analysis for development interventions. Instead it would appear that participatory evaluation practice is not widespread for development organisations in New Zealand, and the design of indicators in a participatory way, even less so. It also seems that the influence of

local and central government through provision of funds and advice has led to the burgeoning use of RBA.

The findings from the local case study organised against the four themes identified in the literature and case examples highlight in particular that:

- Although some clients of Organisation X would like to be involved in all phases of participatory evaluation practice, some would prefer to be involved in only some phases and roles. While they are keen to be involved in the process of developing measures for Organisation X, there is a feeling that there are some tasks and roles that are best performed by staff of the organisation particularly if clients do not have the skills and expertise required to carry out such important tasks;

- the length of time it takes to complete tasks within the four phases is affected by insufficient planning and coordination and impacts on the memory of some participants. Not having a coordinator also affected communication between staff and with clients. The clients did not feel that time had affected their desire to participate in the process. Rather, they despaired that due to a long period of time between the focus group and the time of the interviews, some had forgotten the purpose of the focus groups and what was discussed; and

- there are competing interests between the views held by junior and senior staff, and between staff and clients. While clients were asked to identify Feelings to inform RBA measures, junior staff expressed that measures needed to be useful at operational level, not just at strategic level. How the Feelings would be used as measures was also questioned.

The collective findings from the three research phases are discussed in Chapter Six.

Chapter Six: Discussion

Introduction

An expectation before this study was conducted was that there would be an existing body of literature regarding the application of participatory evaluation principles, and many reported examples of primary stakeholders being involved in practice. Although there is literature that explores the design of sustainability indicators using participatory approaches and how these can be applied, and studies that provide conceptual guidance on involvement of primary stakeholders in participatory evaluations, there appears to be comparatively little written about the practicalities of participatory design of indicators with primary stakeholders for development interventions. Cullen et al.'s (2011) finding that there are few documented cases of participatory evaluation in international development, could account for the limited cases found by this study.

This discussion is, therefore, based on the more limited findings from eight case examples and a local case study. It compares and contrasts the findings of the case examples in Chapter Four and the local case study in Chapter Five. The analysis of these cases is based on the four research questions:

Question 1: At what phases in participatory evaluation practice are primary stakeholders involved?

Question 2: What are the characteristics of outcome indicators designed and used by and with primary stakeholders?

Question 3: What are the dynamics of involving primary stakeholders in design and use of outcome indicators, and evaluation use, and the lessons learned? and

Question 4: What is achieved when primary stakeholders are involved in the design and use of outcome indicators, and evaluation use?

Therefore, the four main headings that align with these research questions that were used in Chapters Four and Five are also used in this chapter. However, the sub-headings used

in this chapter go beyond the scope of those used in the two preceding chapters to reflect the breadth of discussion that emerged from analysis of data.

Phases of involvement in participatory evaluation practice

The data from the case examples suggested that primary stakeholders were engaged at different points in the four phases of involvement in participatory evaluation practice. Using the rubric developed in Chapter Two, it is possible to summarise where most of the activity took place. Figure 6.1 summarises these involvement points by collating the mentions of particular activities in the case examples.

Fig. 6.1: Primary stakeholders' involvement in case examples' participatory evaluation practices
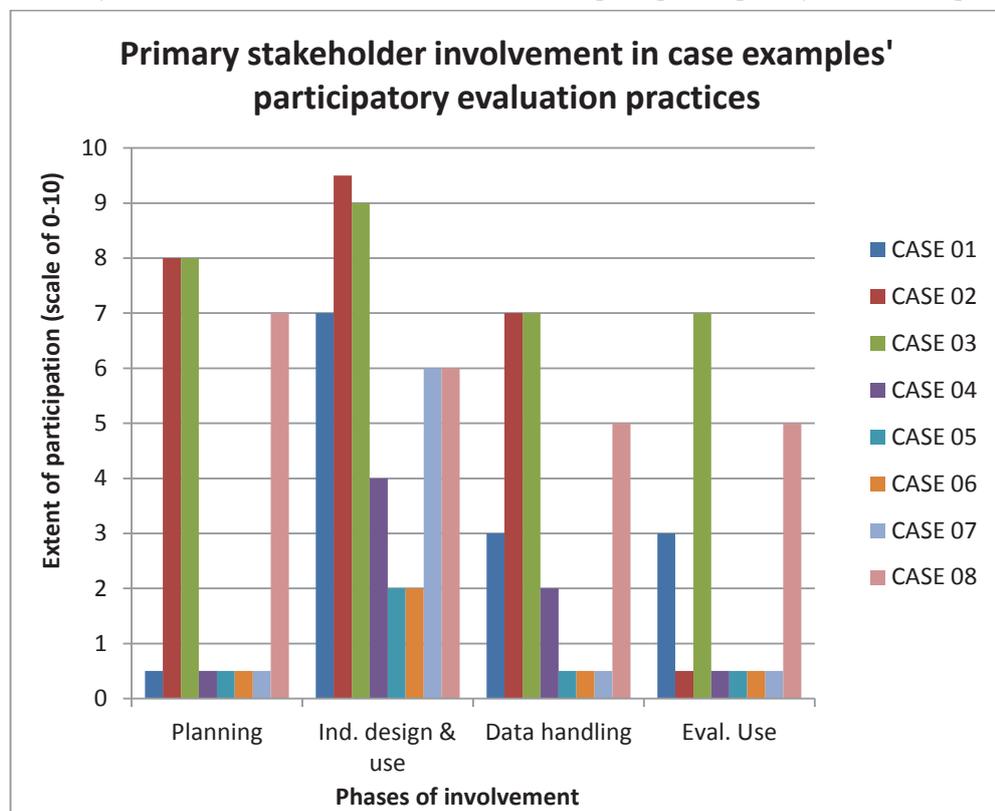
Source: Author.

The graph is based on a 10 point score for each of the phases of involvement with 0 being no participation and 10 being the highest level of involvement. The data presented is based on what was reported, reports which may have been interpreted differently from what the authors meant to convey, and reports that may not have included aspects that are of

interest for this research. Therefore, the scores do not assume '0' or '10' level involvement. Rather, they should be taken on as a visual representation of primary stakeholder participation in the eight case examples.

For example, the increase in participation during the indicator design and use phase was very clear in the case examples due to repeated mention in most cases on the tasks and roles that primary stakeholders were involved in. In other words, all the case examples described participants being involved in the design of indicators in some way and all mentioned at least one task that primary stakeholders were involved for example as data collectors, indicator designers, indicator selectors, indicator aggregators, and/or evaluators.

The local case study has not been included in the analysis in Fig 6.1 because various phases and tasks that Organisation X would like to do in the design of RBA measures have not been completed. To try to put a scoring against each phase would be incomplete and inaccurate, especially when the tasks in the four phases are iterative. It is apparent, however, that at the early stage of Organisation X's process of developing RBA measures clients have not been involved in the planning phase. Their involvement has been limited to the selection of indicators in most cases, and in the case of one client, data collection. Involvement of clients by Organisation X only in the indicator design and use phase is consistent with the findings in most of the case examples reviewed as depicted in Figure 6.1.

One criterion for including case examples in this research was the engagement of primary stakeholders in the task of designing outcome indicators for development interventions. Two reasons appear likely to account for greater involvement of primary stakeholders in the indicator design and use phase. The first is to do with emphasis placed in the case examples on the indicator design task that appears to be seen as central to the evaluation process. The second is primary stakeholder preference to participate in only certain aspects of participatory evaluation practice.

Emphasis on indicator design

One reason that could explain higher engagement of primary stakeholders in the indicator design and use phase is the importance placed on participatory indicator design by staff and facilitators of the eight case examples and the local case study. I could not find evidence of deliberate action on the part of decision makers to exclude primary stakeholders from the other three phases.

Clients of Organisation X, for example, were not involved in the planning phase as part of the steering group. One explanation for this is time pressure. In order to focus on what is contained in the organisation's Business Plan, no time was set aside by the organisation to explore possible involvement of stakeholders in the various phases, tasks and roles when designing RBA measures. It appears the focus is on designing measures and selecting measures to include in RBA, and that the importance of participation in other aspects of the process has been overlooked once the process started. My observation within Organisation X confirmed that while the organisation intends to create a stronger culture of participation, not just in this process but also in other aspects of organisational life, this was not carried through to implementation when designing measures. This finding bears similarity with those of most of the case examples with the exception of Cases 02, 03 and 08. In these three cases primary stakeholders were involved as team members of the group that guided the process. In this way, it seems to me that primary stakeholders were then involved more in other phases throughout the process.

Primary stakeholder preference

Primary stakeholders may prefer to be involved in only some aspects of participatory indicator design rather than in all activities as their time is also scarce. Their interest in participating in any one phase should not be taken to mean that they would prefer comprehensive involvement in all phases, tasks and roles.

Primary stakeholders may want to participate in all aspects of the process because of their belief that they are able to add value to the process not only for the organisation and themselves, but for others who will benefit from the intervention at later stages. The sentiment expressed by two of four clients interviewed is that participation in all aspects results in greater understanding of what the organisation is trying to achieve and, therefore, more meaningful participation by clients, and in clients gaining new knowledge. The words of one client, used as a preface to this thesis, clearly put across the aspirations for further and fuller involvement in participatory evaluation practice (see p. vii). The same point of view is put forward by development evaluation practitioners who state that participatory evaluations work best when stakeholder involvement is in all phases (Cullen et al., 2011). It is not explicitly stated in the case examples whether or not primary stakeholders were given the choice or chance to participate in the aspects that they did, apart from Case 08. What is often reported is the process for involvement of staff and facilitators in participatory evaluation

practice (Cases 01, 04, 05, 06 and 07) in which the evaluation process is instigated by the organisation and staff recruit a facilitator to guide the process.

Secondly, participation in all phases is not always desirable, and does not always result in positive outcomes. Two of the clients interviewed felt that there were some tasks and phases that only staff members of Organisation X should be involved in. Their view was that there may be competing views between the organisation and clients about goals and purposes. By participating in the indicator design and use phase, the primary stakeholders would still have the opportunity to put their views across. Therefore, they accept participating only in some aspects, especially those that they have the competencies to carry out. In Case 08, a similar sentiment was expressed by the Aboriginal land owners who chose to participate only as interpreters during evaluations because that is what they were confident doing (Stacey et al., 2013). Cleaver's (2001) sentiments, that mobilising primary stakeholders to engage in development processes does not always translate to a group of participants who are knowledgeable and have the capacity to contribute sufficiently to produce the desired outcomes, aligns with this view.

The lack of widespread documentation of participant engagement means that greater understanding of processes, phases and positive and negative consequences of participatory evalution are not shared and discussed, (Guijt, 2000) and that lessons learned in case examples like those that are well documented by Estrella et al. (2000) are also not shared.

Characteristics of indicators

The findings about indicators designed by primary stakeholders in the local case study show similarities to the characteristics of those in the eight case examples. The characteristics of the indicators designed by primary stakeholders are depicted in Figure 6.2. These characteristics have been translated into a series of overlapping Venn diagrams that depict the complex interrelationships between them. The goals and objectives of a development intervention have direct bearing on the negotiations that occur and the indicators designed; and negotiations take place when deciding which indicators will be used during an evaluation. The quality of indicators is ascertained through their ability to provide the information required from an evaluation, and this has direct impact on their use. Goals and objectives drive both the development of indicators and the nature and purpose of the negotiations used to secure robust, culturally appropriate, context specific

indicators. Use and quality further determine the simple, visual, quantitative and qualitative nature of the indicators designed.

Figure 6.2: Characteristics of indicators
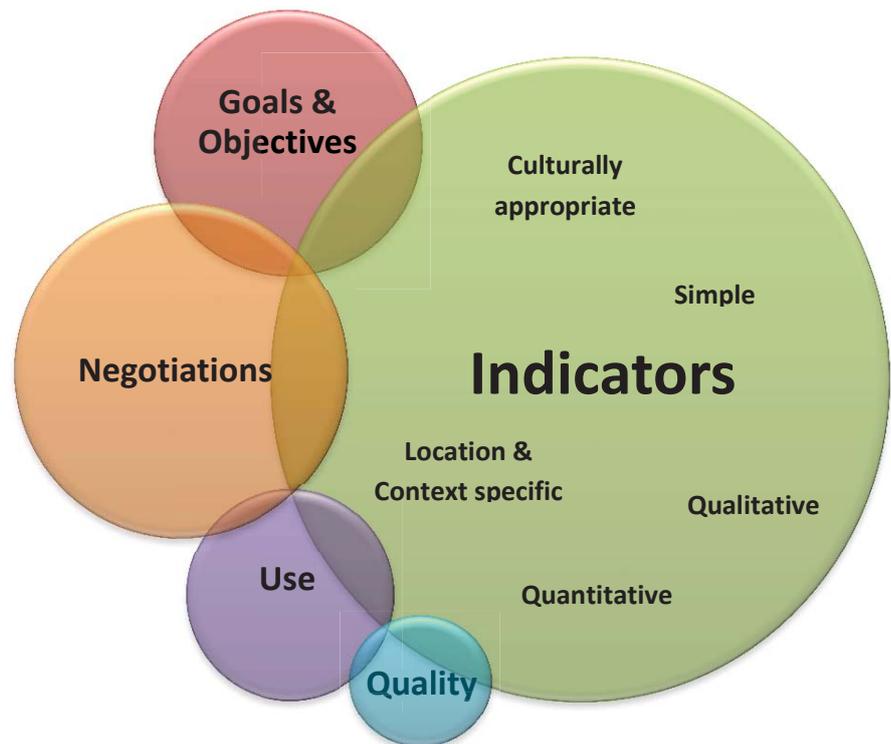
Source: Author

The list of Feelings, for example, collated during the local case study focus groups were linked to the goals and objectives of Organisation X, came out of the discussions around the development of measures, and provided a check for both the quality and use of measures that would more accurately reflect the clients' aspirations. These interconnections are also traceable in Cases 02, 05 and 08.

One difference noted during fieldwork and data analysis was the difference between the negotiations by primary stakeholders in the eight case examples for community

development intervention and those given by the clients of Organisation X for social service interventions. On reflection, it is possible that the negotiations took place in the case examples because all the primary stakeholders benefited, or felt they should benefit, equally from the interventions. They shared some common understanding of the intervention and what they thought it should achieve, and therefore negotiated indicators. For Organisation X, it would seem that because all experiences with domestic violence are personal, there was a level of respect for each other's Feelings that meant they were accepted as presented. Discussions were therefore in support of the Feelings presented rather than negotiation toward the best intervention outcomes.

Based on the sentiments expressed by most research participants, both staff and clients, it is difficult to conclude that the measures developed are simple in the sense that Izurieta et al. (2011) suggest. Unlike Cases 02, 04 and 08 that encourage and give the impression of having designed simple indicators, it would appear that the apparent disconnect between the measures required by the organisation for RBA and the Feelings listed by clients have resulted in some staff and clients stating that the Feelings and the process of fitting them into frameworks are complex but that the statements of Feelings do reflect what most needs to be measured.

It would appear to me that Organisation X's approach to designing measures for its programmes, starting by listing Feelings as part of the process of developing RBA measures, differed from the case examples presented. For instance there is no suggestion that primary stakeholders in the eight case examples designed indicators whose nature was determined by the facilitator or organisation. It would seem that, had the clients talked about the goals and objectives of Organisation X at the start of the focus groups, and then been asked to design measures for the programmes, the measures developed would possibly have been different from the Feelings the organisation sought. This is drawn from the sentiments expressed by clients that include inclination to measure using quantitative data (Interview: C2) and developing measures that are focussed on the services and programmes of the organisation rather than the Feelings (Interview: C4).

What is currently available, the Feelings, are descriptors rather than measures. Descriptors are words or phrases that describe a phenomenon whereas, in the language of RBA, measures track "how well a program, agency or service system is working … and … measures are about the means to get there" (Friedman, 2005, p. 20). The organisation therefore still has steps to take to turn the Feeling descriptors into measures useable in RBA.

Quality of indicators

The framework for performance indicator decision and the performance indicator design template developed by Marr (2009) both offer valuable step-by-step guidelines in the indicator design process (Appendix One). Closely followed, the processes seem to lead to indicators that promise to adequately measure the progress of a development intervention.

A set of indicators was produced in the case examples and Feelings as descriptors for measures for Organisation X. What is apparent is that while satisfaction is expressed in some cases with the indicators designed (Case 01, 02, 03 and 08), there were issues raised relating to: inadequacy of some of the indicators designed where there was insufficient knowledge on design (Case 04); the need for qualitative indicators to be valued more, and for indicators to be culturally appropriate (Case 05); the need for augmenting primary stakeholder indicators with more scientific ones (Case 06); and greater involvement of primary stakeholders in the process in order for them to feel ownership over the indicators and the process (Case 07). In the local case study, the process is not yet complete. There is the desire to fit the Feelings into wellbeing models before deciding how they will become or be used as RBA measures.

The general sentiment put across in the case examples and the local case study is that indicators designed by primary stakeholders reflected their interests. In most cases, the quality of the indicators designed was judged by the relevance to the context for which they are designed rather than an accurate measure for the phenomenon to be evaluated. This makes it difficult for those wanting to assess how robust the indicators are.

Fitting Feelings into frameworks

The decision to fit the Feelings into Te Whare Tapa Wha and/or Five Ways to Wellbeing models is possibly achievable. However, doing that, and then placing these models into RBA framework is likely to be complex. Two of the case examples (Case 06 and 08) all took the indicators designed by primary stakeholders and fitted them into existing frameworks. This appears to have been successful because the frameworks and the measures contained were aligned to the work being done by the organisations the indicators were being designed for.

Concerns were expressed by Organisation X staff during interviews about the usefulness of fitting the Feelings into Te Whare Tapa Wha and Five Ways to Wellbeing, and then trying to fit existing frameworks into yet another form. In their current form, it is difficult to see how the Feeling descriptors, meant to be used as measures for the third RBA question,

will align with the draft RBA measures that have been developed for the first two questions, and what is already contained within the third question as placeholders (See Figure 5.3).

Looking at the Feelings, the RBA measures designed by the organisation, and the aspirations of Te Whare Tapa Wha and Five Ways to Wellbeing it may be difficult and unnecessary to fit the Feeling descriptors into one model. What may be possible for Organisation X is to create two products from this process. The first could be presentation of the Feelings incorporated into Te Whare Tapa Wha and Five Ways to Wellbeing as aspirational models that are discussed with clients in educational groups or other work streams. The second product could then be an RBA scorecard that incorporates the measures currently in place for the first two questions of RBA, and the measures for determining the difference the organisation has made. This process could still be participatory and involve clients, and focus specially on performance in line with RBA, a desire expressed by coalface staff. My impression from observation and discussions is that the process is being over-complicated and may result in being 'arduous', like in Case 04.

Dynamics of involvement and lessons learned

During the process of reviewing the eight case examples, and during fieldwork, it became apparent that there would always be consequences of having more than one stakeholder group involved in participatory evaluation practice, especially in relation to indicator design. Having skilled facilitators; cultural awareness and competence; understanding the nature of participation by primary stakeholders and how this was affected by power; and time scarcity, were important influences on the experiences reported. What is presented in Chapter Four, in reviews of the case examples, and Chapter Five in the local case study, is not revisited here. Rather this section looks briefly at these aspects and in greater depth at power dynamics and time scarcity.

Facilitation

Facilitation skills of those in the local case study varied, as did those in the eight case examples. While Case 04 reported on the effects of not having adequately skilled facilitators on the process of designing indicators, lack of prior training in facilitating the focus groups in the local case study does not appear to have impacted on the quality of discussion, or the outputs of each one. Instead, it appears that some staff and one client of Organisation X

learned from drawing on skills they have in facilitating other processes, and in the one case of client C1, observing and learning alongside a skilled facilitator.

Culture

The importance of honouring cultural practices is seen in the local case study and in Case 08. Both Case 08 and the local case study observed cultural practices in ways that were relevant in their situations. For instance, in Case 08, processes were delayed to allow time for Aboriginal land owners to attend funerals, while in the local case study, the simple act of starting a focus group by sharing food with Māori clients aligned with cultural practice.

Participation

Participation by stakeholders in any process is affected by power dynamics between and within stakeholder groups in the local case study. Because the process of developing measures for RBA involved staff and clients separately in the local case study, with interaction between the two groups limited to facilitation of focus groups, power dynamics played out largely within staff and steering group interactions, and between facilitators and clients.

Although some clients expressed that they were comfortable in the focus groups and did not feel any pressure from staff facilitators to participate, this was not the feeling of all clients. The settings, facilitators, number of participants, and the experiences of participants at particular points in their lives resulted in different experiences of the focus groups. I observed that staff made every effort to ensure that clients participated of their own volition and in an environment that was conducive for open and honest talk. However, power dynamics were noted in the focus groups, and in the steering group setting. Three points are raised here: the importance of seeking ways to minimise the effects of unequal power relations in advance, recognising the competing interests of different stakeholders and seeking ways to accommodate different viewpoints; and the need to carefully assess the usefulness of multi-stakeholder participation.

Minimising power dynamics

Although there was some effort on the part of local case study staff members to lessen the possible effects of power on the focus groups, it may have been better if the focus groups had been facilitated by coalface staff. Most staff members who were interviewed expressed concern about the involvement of senior management in the focus groups. Apart from coalface staff gaining valuable skills in the process, the view expressed was that, because they

work with clients in the different programmes, they would have been able to draw on their skills in relating to these clients on a day-to-day basis in the focus group setting. As it was, one client felt uncomfortable because she regarded participation of senior staff as being unfair for a fellow client (see Chapter Five). Mitigating such power dynamics by involving coalface staff may need to be tested in further research, but seems a plausible argument. Conversely, however, involving senior staff members may have signalled to the clients that their views were being taken seriously. This could also result in discussions on the competing views between strategic and operational interests being raised early on in the process.

Another idea would be to have an external facilitator, knowledgeable of family violence, to work with staff and clients in the process. In Case 07, as with the local case study, staff facilitated the process and Parkinson (2009) reports extensively on the power dynamics between staff and the farmers who were primary stakeholders in that case. According to Parkinson (2009) because participatory monitoring and evaluation was introduced by programme managers, farmers felt that it was for staff benefit rather than their empowerment, and participated as a form of patronage. This sentiment was not expressed in the local case study. In Cases 01 and 08 external facilitators who appear to have been experienced in the particular field of the interventions for which indicators were being created, worked with different stakeholder groups to create outcome indicators. This appears to have mitigated against the expression of unequal power dynamics between staff and primary stakeholders, particularly in Case 08.

Strategic versus operational interests

Hamilton et al. (2000) refer to competing interests amongst and within stakeholder groups that emerge during discussions about indicator design. They state that negotiations are required to ensure that indicators designed accommodate differing views. In the local case study, the tension between the two types of indicators, strategic versus operational, is seen in the lack of understanding expressed by coalface staff members about how the Feelings will be useful in measuring the effectiveness of programmes and services offered by the organisation. Clear and on-going communication and discussion between junior and senior staff can alleviate some of these tensions. It would seem from a researcher perspective that based on the aspects of work that the RBA measures are being designed for, operational rather than strategic measures are required, and that involvement of junior staff could be instrumental in working out with clients what these might include. The same issues surfaced in Case 07 where

the strategic goals of senior management appeared to not have been shared by staff working in the field.

Length of time

It has frequently been pointed out that participatory evaluation processes take longer to complete than non-participatory approaches. The impact time has to date had on Organisation X's process raised some interesting perspectives. At the time of interviewing clients of Organisations X, all of them, except one who is now a staff member of the organisation, had not been involved in any aspect of the process or heard anything about its progress for between 10 and 12 months. The general feeling by both clients and staff was that although the process was taking long, it was important and needed to be given the time it needed to allow things to be done correctly. Negativity was expressed more in relation to the effects of time on memory; ability of Organisation X to keep all clients who took part in the first focus groups involved in future processes; and loss of momentum.

The length of time it takes to go through participatory processes is cited often in evaluation. In Case 04 and 08 the amount of time it took for different stakeholders to engage with and participate in processes, and the effects of such time, is highlighted. The impact of unplanned breaks (Izurieta et al., 2011) and insufficient facilitation skills (Hillyer & Purohit, 2007) can have on participatory evaluation practice can affect the quality of outcomes. In the local case study a cause of lengthy time lapses was attributed to lack of coordination and consequent lack of clarity about what was to happen next. The effects of time lapse were on the memory of the clients to recall clearly earlier focus group discussions.

Coordination and communication

Lack of clarity about who was steering the process and what the next steps were to be, contributed to loss of momentum and some confusion during the process in the local case study. My experience as researcher with Organisation X is that lack of adequate communication amongst steering group members has contributed to this. For instance, the lack of understanding and clarity about how various aspects of the process were linked and would unfold, as communicated to me by staff, resulted in lapses in time that may not have occurred had there been better communication. This led to the view that the process was disjointed (S1, S2) with staff thinking that there were probably a lot of meetings they had missed, when in fact that was not the case. Lack of communication on the direction of the

process appeared to have arisen partly because all those involved in the process were learning from the experience as it unfolded, and were invited to participate at different points in the process. In *Learning from Change: Issues and experiences in participatory monitoring and evaluation* (Estrella et al., 2000) at least half the case studies pointed out the need for improved communication between and within stakeholder groups.

There was great willingness on the part of staff and senior management to ensure that the process did not seem tokenistic because of time constraints. The staff suggestion to have a coordinator could be one way of resolving the time challenges currently experienced. Because staff have work pressures and, at times targets that need to be met, approaching volunteers to assist in the coordination of such processes from organisations such as the local Volunteering City (e.g. Volunteering Auckland) would be valuable. In the case of Organisation X, they already have a volunteer consultant who is willing to assist. It may be that the facilitator of the process could take on the coordination role to ensure that poor communication does not negatively affect the outcomes of the process.

What I have noted from the case examples is that when there is a researcher, evaluator or facilitator coordinating the process, as it was in Cases 01, 02, 03, and 08, there is likely to be less time taken in completing tasks due to deadlines and timeframes. In the case of this research, I wanted as much as possible to not affect the process and the length of time it took to develop RBA measures, and I was not leading the process in any way. In my discussions with Organisation X I communicated that I would rather the organisation take as long as they need to in their process, than be rushed into aspects they are not ready for in order to fulfil my research requirements.

Temporal effects on memory and participation

When interviewing clients, my observation was that they struggled to remember the content of the focus group they had participated in, and more often than not asked to be reminded of what had transpired. Clients communicated that time had affected their ability to participate in my research with clarity about the focus groups they had participated in, and they also felt that as more time passed participation in the next phases of Organisation X's process would be affected. It was intimated that changes in their lives since they had been involved in the first focus group would affect their contributions. When setting up interviews one of the clients who participated in a focus group was willing to take part in this research but had moved to a city more than 450 km from the location of Organisation X. For safety

reasons, she would not be returning to the location of this study. When taking the next steps after there have been long lapses of time, it would seem important that Organisation X spends a significant amount of time talking about the focus groups and the Feelings, and the phases and tasks so that there is shared understanding on what people are trying to achieve. Such issues were not raised in the eight international case examples reviewed.

Time as a negative consequence of participatory evaluation has been commonly talked about (Cullen et al., 2011; Estrella & Gaventa, 1998). (Cullen et al., 2011) found that, in some cases, time becomes a barrier that results in evaluations using traditional non-participatory evaluation methods.

Achievements

Although Organisation X's process of creating RBA measures was not complete at the time of interviewing participants, there is a real sense on the part of the clients that the process was empowering, largely because they had been asked to participate, and because they felt heard.

It was clear in my discussions with the clients that they felt valued because they had been asked to be part of what they considered to be a very important process. In the case of two clients it was clear that the experience of the focus group had impacted on their self-esteem, in that they talked about what they were able to learn of themselves and how the focus group had contributed to their growth. The feeling that clients had given back to the organisation by way of their participation is also noteworthy. Reciprocity through this process can also be seen as empowering as the clients felt that they had something of value, their Feelings, to contribute to the process. Achievements such as this, apart from being noted by a researcher or facilitator, may not be made known to an organisation unless there is clear review and documentation of experiences.

In the eight case examples reviewed, the feeling that primary stakeholders were valued and valuable in participatory evaluation practice also came through. Based on Table 4.3 it was clear that empowerment was achieved to varying degrees, dependent on the phases and roles of involvement of primary and local stakeholders. The three case examples (Cases 02, 03 and 08) that had high involvement of primary stakeholders in the four phases of

participatory evaluation practice of planning, indicator design and use, data handling and evaluation use, appear to have recorded greater empowerment of primary stakeholders.

Guijt (2000, p. 204) points out that "for indicator development to be empowering is an impressive feat and one that few monitoring and evaluation efforts can correctly claim to have achieved". A theme that emerges from the literature, when talking to staff and clients especially, was the possibility of critical consciousness in such processes. Bourdieu and Wacquant (1992) speak of critical consciousness as being the prospect of a person's usual ways of being, such as thinking and acting within a certain cultural way, being affected when they are socialised to new ways of life and perceptions. Three of the clients I spoke to talked positively about the experience in the focus groups and how possible involvement in other aspects of participatory evaluation processes could lead to such critical consciousness. Their view was that involving clients in more aspects of evaluation practice could incentivise clients to pursue careers or involve themselves in similar work. One client in particular was motivated enough by her reflection on growth during and after the focus group to consider further study.

Conclusions

The aim of this thesis was to examine the involvement of primary stakeholders in the design and use of outcome indicators and evaluation use for development interventions, with specific focus on the phases that primary stakeholders are involved in; the characteristics of indicators designed; the dynamics of involving primary stakeholders and the achievements that come out of such practices.

While the literature review on participatory monitoring and evaluation conducted by Estrella and Gaventa (1998), and the practical aspects of participatory monitoring and evaluation suggested by Guijt (2000) provided a starting point for the arguments of this thesis, an analysis of international literature and eight international case examples indicated a need for more systematic assessment and use of more holistic criteria. The assessment undertaken alongside analysis of local case study material, has allowed me to broaden the discussions on the involvement of primary stakeholders in indicator design processes and to offer a deeper understanding of the practical considerations such involvement entails.

The development of a rubric for assessing stakeholder involvement in participatory evaluation practice was a primary tool for my analysis. The four phases identified were planning, indicator design and use, data collation, and evaluation use. The eight case examples reviewed in Chapter Four and the local case study in Chapter Five tell stories of varying degrees of inclusiveness of the perceptions and knowledge that primary stakeholders hold. It is evident that primary stakeholders are more involved in the design and use phase than any other when the focus is on creating outcome indicators for an intervention. In participatory contexts, there appear to be no deliberate effort on the part of decision makers to exclude primary stakeholders from aspects of participatory evaluation processes. Rather it would seem that attempts by decision makers to make the process work well for primary stakeholders may get in the way of having them around the table during discussions on the different tasks within the four phases.

More research needs to be done to further understand why primary stakeholders are not routinely involved in the planning, data handling and evaluation use phases. The importance placed on preparing for primary stakeholders to be involved in designing indicators appears to overshadow the opportunities for involving them in all phases.

Whether involvement of primary stakeholders in other phases would have positive effect on the characteristics of the indicators designed is a moot point. What is clear is that when primary stakeholders are participants in the four phases they are more likely to understand and value the indicator design and use (Holte-McKenzie et al., 2006; Izurieta et al., 2011; Vernooy et al., 2006). The data has shed light on the first research question related to phases of involvement, and determined that primary stakeholders tend to have limited involvement in all but the indicator design and use phase.

The second research question examined the characteristics of the indicators designed by primary stakeholders. The data has revealed that these indicators are often linked to the goals and objectives of the organisation or the particular intervention for which indicators are designed. The indicators designed are said to be culturally appropriate and contextually relevant for particular development interventions. A visual presentation that depicts the relationship between the different characteristics of indicators emerged in the analysis of the case examples and the local case study.

An important issue raised in the literature is that linking indicators to the goals and objectives of an organisation or a development intervention makes it more likely that change can be measured in ways that are more meaningful to the programme participants. Negotiations during the indicator design and use phase facilitate clarification around competing interests and views. The skills of an experienced facilitator are identified as being useful for ensuring that the different perspectives and views on the indicators are catered for.

When one stakeholder group is involved in the design of indicators, as was the case with the local case study and with Case 01 the negotiations and/or discussions take place within the stakeholder group. It was found that negotiations may take place between two or more stakeholder groups and will involve negotiations because of competing interests, often between primary stakeholders and staff (Izurieta et al., 2011). Such negotiations are important because the indicators selected often reflect the interests of stakeholders.

The evidence points to resourcefulness and agency on the part of primary stakeholders who usually suggest indicators steeped in the reality within which they live and experience. The quality of the indicators is often not measured against existing criteria such as CREAM and SPICED (Roche, 1999) making it difficult for those wanting to systematically assess the robustness of the indicators to do so.

An aspiration of this research was also to begin to understand the dynamics of involving primary stakeholders in the design of indicators in particular. In response to the third research question, a conclusion drawn is that there are both positive and negative consequences for involving primary stakeholders. The research confirmed the prior expectations I held based on insights from the literature and my own experiences on the need for good facilitation skills, cultural awareness and participation. I expected that time would be cited as a negative impact in participation evaluation practice.

While it would appear that in most of the case examples and in the local case study that the facilitators had the skills to take primary stakeholders through the process of designing outcome indicators, the research shows that lack of adequate skills can adversely impact on the experience for both the facilitator and stakeholders. Cultural awareness was seen to be important for primary stakeholders to ensure that they are able to participate in evaluation practice, while observing cultural practices. As shown through the local case study and the case examples power dynamics are felt and interpreted differently by primary

stakeholders. A conclusion drawn is these can be minimised through open, reflective, dialogic processes.

Time constraints appeared as a significant issue in discussions in the local case study. The length of time it had taken for the organisation to communicate with clients had been deemed too long. A conclusion reached was that a coordinator responsible for communication and keeping within timeframes can mitigate concerns about both length of time and quality and frequency of communications.

The last research question focussed on finding out what is achieved when primary stakeholders are involved in the design of outcome indicators and participatory evaluation practice. To varying degrees, primary stakeholders have, in the very least, found some spaces where their views count, and in some cases have experienced joint learning and had their skills developed. In the local case study the clients felt empowered by simply being asked to be the ones to develop measures for RBA. The finding that appropriate plans and indicators are devised by primary stakeholders confirms that with adequate facilitation skills development and guidance, primary stakeholders are capable of producing quality outcomes.

Thus while current practice in much participatory development practice limits the role of primary stakeholders to one phase of evaluation practice, there is potential for this limited view to be challenged and practice extended to all phases. Further research across a range of different real world case studies could reveal where opportunities exist for such expansion.

Limitations of this study

I was unable to work with an organisation that was conducting participatory evaluation or designing outcome indicators for a development intervention. Instead I worked with an organisation in New Zealand that supports individuals and families to become violence-free. While the methodological considerations in this case study align to those for development interventions, sentiments expressed by clients were often about individual aspirations borne out of negative experiences in life, rather than sentiments that might be expressed by a community working together to overcome family violence. The focus in the thesis on the methodological considerations was an important framing for understanding practice both in the New Zealand context and across the eight international case examples.

Secondary data was relied on to understand at what phases in evaluative processes are primary stakeholders involved in the design and use of outcome indicators and evaluation use, and while the data used from the eight case examples was valuable for informing this study, it was not without its limitations. As is the case with all secondary data, the literature obtained on the case examples was not created for the purposes of this study. I had to understand it within the context it was written and interpret the data for this research. This entailed being careful not to misrepresent data in a manner that was not intended. The data is also different for each of the case examples and not always easy to compare. Thus qualitative thematic analysis was used to draw out the key issues in participatory evaluation practice.

Closing

It is apparent that, under the banner of participatory evaluation practice, primary stakeholders have taken part in the process of designing outcome indicators using varying approaches, frameworks and methods. However, it is also apparent from the local case study discussions that, in New Zealand at least, there is little recourse to generally accessible knowledge about indicator design. What appears to be more the case is that individual organisations might work in more ad hoc ways 'inventing as they go along'. This might further suggest there is a need for more systematic capability building across the NGO sector so that the ideas and frameworks of participatory evaluation and indicator design are better understood, both at theoretical and practical levels. More local research could usefully extend our knowledge of how frequently NGOs are working in isolation and reinventing ideas that are more systematically understood and explained in local and international literature.

Appendix One: Performance indicator templates

Performance indicator design framework

1. Which strategic element do we want to assess?
2. Do we have a key performance question to answer?
   Yes / No → Don't measure, rethink!
3. Are there decisions a potential indicator would support?
   Yes / No
4.1 Can we use existing methods for the potential indicator?
   No / Yes
4.2 Design new assessment method

- Indicator name
- Measurement instrument?
- Source of data?
- Metric/formula?
- Collection date and frequency?
- Who measures?
- Expiry date/revision data?
- Targets/benchmarks?
- Costs of indicator?
- Audience?
- Reporting frequency?
- Reporting channels?

5. Can we collect meaningful data for this indicator?
   Yes / No
6. Does the indicator help us answer our KPQ?
   Yes / No
7. Does the indicator help us make better decisions?
   Yes / No
8. Does this indicator invite cheating?
   No / Yes
9. Are the assessment costs and efforts justified?
   Yes / No
10. Start collecting the indicator

Source: Marr, B. (2009). *Managing and delivering performance*. Amsterdam, The Netherlands: Butterworth-Heinemann.

Performance indicator design template

Performance Indicator Design Template	
The Basics – Why do we need this indicator?	
1. Indicator name	Provide a clear indicator name.
2. Strategic element being assessed	State which strategic element of objective this indicator relates to.
3. Owner of strategic element	Identify the person(s) or function(s) responsible for the strategic element or objective.
4. Key performance question	Provide the key performance question this indicator is helping to answer.
5. Decisions supported	List the decision(s) this indicator is helping to support.
How will the data be collected?	
6. Data collection method/ Instrument	Describe the measurement instrument used and how the data is being collected.
7. Source of data	Identify where the data for this indicator comes from.
8. Formula/Scale/Assessment	Explain the scale or formula used to assess performance.
9. Frequency	Illustrate how often the indicator is measured.
10. Who measures/ Reviews the data?	Name the person who is collecting, updating and/or reviewing the data.
11. Expiry/Revision date	Identify until when this indicator will be collected or when it will be revised.
What are the targets?	
12. Targets/Performance thresholds	Set the targets and/or benchmarks for this indicator and provide performance thresholds (e.g. for when traffic lights turn from green to amber to red, and so on).
How good is the indicator?	
13. Cost Estimate	Estimate of the costs incurred by introducing and maintaining this indicator.
14. Confidence level	Provide an evaluation of how well this indicator is measuring what it is supposed to measure (e.g. written comment/evaluation: e.g. good☺ fair☺imperfect ☺).
15. Possible dysfunctions	Note down any possible ways this indicator could encourage cheating.
Who will see the data? How will the data be presented?	
16. Audience/Access	Identify the audience of the indicator data and outline who has access rights to it.
17. Reporting frequency	Illustrate when and how often the indicator is reported.
18. Reporting channels	Describe the channels used to report this indicator (reports, meetings, online, etc.).
19. Reporting formats	State how the performance data is presented (numerical, graphical, narrative formats)
20. Notifications/Workflows	Identify any notifications, e-mail alerts and workflows triggered by this indicator

Source: Marr, B. (2009). *Managing and delivering performance.* Amsterdam, The Netherlands: Butterworth-Heinemann.

Appendix Two: Informed Consent

Participant Consent Form

This informed consent form is for clients/staff of Organisation X who I am inviting to participate in my Masters of International Development research entitled "The dynamics of involving primary stakeholders in designing outcome indicators and evaluation plans, and evaluation use."

Please note that my research proposal was reviewed by the Development Studies Programme and the Ethics Committee at Massey University. I was interviewed and applied to do this research with Organisation X. Both the Development Studies Programme and the Ethics Committee at Massey University are in agreement that those who participate are protected from harm. Both have given the go ahead for me to do this research with those who would like to participate voluntarily.

Position:	Student Researcher
Name:	Shupayi Mpunga
Institution:	Massey University
Programme:	Master of International Development
Project title:	The dynamics of involving primary stakeholders in the design of outcome indicators and evaluation plans, and in evaluation use.
Supervisors:	Dr. G. Prinsen
	Lecturer, School of People, Environment and Planning
	and
	Associate Professor R. Peace
	School of People, Environment and Planning

This Informed Consent Form has two parts:
- Information Sheet – so that you know what this research is about; and
- Certificate of Consent – for you to sign if you agree to participate.

Information Sheet

Introduction

Kia ora,

My name is Shupayi Mpunga and I am a student at Massey University studying towards a Master of International Development. I am really interested in learning about outcome measures/indicators that are designed to determine the value, merit and worth of programmes. I would like to give you information and invite you to be part of my research. Before you decide please feel free to talk to me or Organisation X staff you feel comfortable with about the research.

There are words and phrases I have used in this form that you may not understand. Please ask me about these and I will explain as best I can. You can also ask me questions if you do not understand any part of my research.

Purpose of the research

I would like to find out what happens when people who benefit from a programme are involved in deciding what outcome measures will be used to judge the value, merit and worth of a programme, and how the same people will be involved in evaluations and use of the evaluation findings. Basically I am looking at your participation in designing outcome indicators and evaluation plans. I would like you to talk to me about your views on Organisation X's evaluation planning process and your involvement because people have different views about who should participate, at what level, and in what way.

Research procedures, confidentiality and anonymity

This research will involve you taking part in an interview that will last about one hour (or as long as you would like to talk to me) and a group discussion that will take about an hour and a half. I would like to hear about your experience and thoughts about the process.

Procedures - During the interview or group discussion I will start by introducing myself and answer any questions you may have about my research and the interview or discussion. I would like to hear your thoughts so will ask very few questions so that you can tell me what is important to you. This will be a time to reflect on the process and share your views. You do not have to answer all the questions I ask and you can talk about any part of the process that you want to. The interview will be just you and I and will take about an hour. The group discussion will be with other Organisation X clients. I will ask a few questions that you can all chat about. It should take about an hour and a half.

Confidentiality and anonymity - No one will know what you have said in an interview apart from me. It will all be confidential and anonymous. If you do not want it known that you took part in an interview I can meet you outside of Organisation X offices or at a place you are comfortable with. You will not be asked to share any personal experiences. The interview will take about an hour – could be shorter or longer depending on what you are comfortable with.

Like the interview the discussion will only be about the evaluation planning process and there will be no questions about your personal life. Confidentiality cannot be guaranteed in a group setting but I will be asking everyone who participates to respect everyone's views and keep them to the confines of the group discussion.

Voluntary participation and withdrawal

You do not have to take part in this research. Participation is on voluntary basis. Whether you choose to participate or not, the services you receive through Organisation X will remain the same. If you agree to participate and then change your mind at any point during the research process, please let me know and I will respect your wish to withdraw from the research and will not use any of the information for my thesis.

Risks

You will not be asked to share any personal information about your life and your personal beliefs. This research is about Organisation X's evaluation process. However if you feel uncomfortable answering any questions then you don't have to answer. There is also a chance that you may share some personal or confidential information by chance. If you need support during or after the interview, staff at Organisation X will be able to help out.

Benefits

There will be no direct benefits to you. Your involvement however will be helpful in providing me with information to put together a thesis that organisations wanting to go through a similar process can access and learn from. There will be no monetary incentive to take part in the research.

Sharing the Results

Your name will not be mentioned in the results at all. The information I get from the interview will be coded. You will get a chance to review the information and tell me if I understood your views before I put them together with other people's views. Information from group discussions will be shared with those who take part again for verification. Once I have a draft report I will share it with you and other Organisation X stakeholders. You will receive a summary of the results at the end.

Contact details

If you have any questions please ask. You can contact me on:

021 0622481 – you can text and I will call you.

360 2362 in the evenings or shupayi.mpunga@gmail.com

Certificate of Consent

1. I have been asked to participate in the research on "the dynamics of involving primary stakeholders in the design of outcome indicators and evaluation plans, and evaluation use". I have read the information about it and/or it has been explained to me. I was given the opportunity to ask questions before the research took place. I will be participating in this research voluntarily.

2. It has been explained to me that I am free to withdraw from the study at any time, or to decline to answer any particular questions in the study.

3. I agree to provide information to the researchers under the conditions of confidentiality and anonymity set out on the information sheet. It has however been explained that during group discussions participants will be asked to respect each other by keeping discussions confidential and that confidentiality during these group discussions will not be guaranteed.

4. I wish to participate in this study under the conditions set out in the Information Sheet and voluntarily.

5. I would like the information I provide to be discussed with me before it is used for any research purposes.

6. It has been explained that the information collected for the purposes of this research study may be used at conferences or for other academic purposes. I agree to this being done.

Participant's Name: _____ **Signature:** _____

Contact details: _____

_____

I confirm that I have explained to the participant what this research is about and have given him/her the opportunity to ask questions which I have answered honestly. I have in no way coerced the participant to give consent which has been given freely and voluntarily.

I have given a copy of the form to the participant.

Researcher's Name: _____ **Signature:** _____

Contact details: _____

Date: _____

Appendix Three: Interview Questions

Semi-structured interview questions - staff[8]

1. First thoughts best thoughts! Tell me what you think and feel about the process so far of involving clients in the development of RBA measures.

2. What has worked well and what hasn't?

Design of measures

3. To what extent do the clients' Feelings align with the goals of Organisation X?

4. The clients were asked about their feelings past, present and future. How do you think these will be turned into measures to answer the third RBA question: What difference did it make?

5. The idea of putting the Feelings into frameworks – Te Whare Tapa Wha and Five Ways to Wellbeing – was mooted. Why and how do you think this will align with RBA?

6. One criticism levelled against participatory indicator design is that the indicators are often specific to local context. They often cannot be generalised and used for other interventions. Any thoughts on this in relation to the measures being designed in this process?

Phases of participation and roles

7. (Rubric for assessment of stakeholder involvement was explained to participants).

   In future, for this process or others, where would you like to see clients involved in the four phases of evaluation practice?

   What are future benefits and opportunities, and possible challenges and barriers?

---

[8] The questions were modified for each staff interview depending on the role each has had in the process of developing RBA measures for Organisation X.

Dynamics of Involvement

8. **Facilitation** – Four focus groups were held. Four facilitators were involved in these. What do you think went well and what didn't?

   o Had you facilitated such a process before? How confident were in facilitating the focus group? What did you learn?

9. **Culture** – what cultural considerations were you aware of in the process?

10. **Participation** – Think about the focus groups and general participation by clients. Can you comment on participation:

    o The nature of that participation (active/passive).
    o Can you think of anything that may have enhanced the quality of participation?
    o Numbers of those who chose to participate.

11. **Power dynamics** in evaluation can be between the researcher/evaluation and staff, staff and clients of Organisation X, consultants, etc. but also within the same stakeholder group. Can you comment on this aspect in this process.

    o Relationships amongst stakeholders
    o Age and education
    o Social standing
    o Any other

12. **Time** – In terms of time, what do you think of the time it has taken to get to the point the process is at the moment?

    o Any effects this may have on future aspects.

13. **Cost** – what do you think the cost of involving clients was vs not having them as part of the process?

Achievements

14. What do you think the benefits of being involved have been for clients?

15. How has the organisation benefitted from client involvement in the process so far?

16. What are next steps? Or What do you think the next steps should be?

Semi-structured interview questions - clients[9]

1. First thoughts best thoughts! Tell me what being involved in the design of measures was like for you.

2. What were some highlights from that focus group, and what were lowlights?

Design of measures

3. What is your understanding of what indicators and evaluations are?

4. The facilitator explained to you that Organisation X wanted to design indicators of success from a client perspective. What did you understand about this?

5. You were asked about what you would like to feel in future. Based what I have understood and explained about what Organisation X would like to achieve through this process, what do you think needs to be done next with clients?

6. There is an idea of putting the Feelings you spoke about into frameworks such as Te Whare Tapa Wha and 5 Ways to Wellbeing. Do you know anything about these? If so, can you comment on this?

7. From your understanding what are Organisation X's goals and objectives and what do you think the links are to the Feelings?

8. You are designing the indicators as part of RBA. Do you know what RBA is?

   o Can you comment on how you think the Feelings could inform Organisation X's RBA measures?

Phases of participation and roles

9. (Rubric for assessment of stakeholder involvement was explained to participants).

   In future, for this process or others, which phases, tasks and roles would you like to be involved in?

   What are the benefits and opportunities, and possible challenges and barriers?

---

[9] Start by reminding clients of the focus group they were involved in and the Feelings that they listed.

Dynamics of Involvement

10. **Facilitation** – there were 4 focus groups. These were facilitated by Organisation X staff. How comfortable were you with the way your focus group was facilitated?

11. **Culture** – what cultural considerations were you aware of in the process?

12. **Participation** – Think about the focus groups and general participation by clients in the organisation, can you comment on participation:

    o What are the current spaces for participation in Organisation X as an organisation?
    o Numbers of those who chose to participate. (How many people were doing the peer support training?)
    o The nature of that participation (active/passive).
    o Were any incentives given for participation?
    o Can you think of anything that may have enhanced the quality of participation?

13. **Power dynamics** in evaluation can be between the researcher/evaluation and staff, staff and clients of Organisation X, consultants, etc. but also within the same stakeholder group. Can you comment on this aspect in this process.

    o Relationships amongst stakeholders
    o Age and education
    o Social standing
    o Any other

14. **Time** – In terms of time, what do you think of the time it has taken to get to the point the process is at the moment?

    Any effects this may have on future aspects.

What was achieved?

15. From your perspective as a client/volunteer what did this participation mean for/to you? What was the importance/significance?

    o Appropriate indicators?

16. What do you hope will happen next in terms of your involvement?

Appendix Four: Rubric

## Rubric for assessing involvement of stakeholders in evaluation practice

Phase	Tasks	Explanation	Role	Extent Measure		
				Not really (0-3)	Partially (4-6)	Fully (7-10)
1 - Planning	Decision to design indicators or evaluate	The initiation of the process to design outcome indicators or conduct evaluation.	Initiators			
	Guide evaluation process	The manner in which the process of formulating outcome indicators or evaluation plan is guided and led.	Evaluation team members			
	Select evaluation framework	The approach/framework/model used to design the research/evaluation.	Evaluation designers			
	Needs assessment	Initial process of finding out what needs to be done e.g. participatory evaluation.	Needs assessors			
	Capacity building	Skills development of various aspects of participatory evaluation practice.	Trainers			
2 - Indicator Design & Use	Data collection	The process of gathering data to design indicators or evaluate.	Data collectors			
	Design of indicators	The exercise of designing outcome indicators.	Indicator designers			
	Selection of Indicators	The exercise of selecting or prioritising the indicators to be included in the monitoring and evaluation plan.	Indicators selectors			
	Aggregation of indicators	The process of aggregating two or more sets of indicators usually for comparison across a number of sites e.g. villages.	Indicator aggregators			
	Evaluation	The exercise of determining value, merit and worth of an intervention.	Evaluators			
3 - Data Handling	Data analysis	Analysis of data collected during indicator design phase or evaluation.	Data analysts			
	Reporting	Decision on what goes into the evaluation plan; and/or the evaluation report.	Report writers			
	Dissemination of Information	The distribution of information on outcome indicators included in the evaluation plan and/or the evaluation findings.	Information distributors			
4 – Evaluation Use	Implementing of findings	Decisions about how evaluation will be used to improve the outcomes of an intervention, and then making those changes.	Evaluation users			

Appendix Five: Wellbeing models

Te Whare Tapa Wha

Te Whare Tapa Wha is a holistic health model that was developed by Māori health workers in 1982 (Rochford, 2004). It recognises the four cornerstones of health as being taha wairua (spiritual), the pertinence of identity in relation to culture and connections with others, ancestors, and the natural environment; taha hinekaro (emotion), Māori ways of feeling, thinking and conducting oneself; taha tinana (physical), bodily health; and taha whanau (Ministry of Social Development), social wellbeing (Durie, 2011). Articulation of this Māori health model by Māori resulted in changes in mental health care that saw greater participation by Māori leadership and health workers at administrative, clinical, and governance levels (Durie, 2004). In order for Māori to be well, Durie (2004) asserts that their spiritual, psychological, cultural and economic aspects of life have to be developed. Figure A is an adaptation of Rochford's (2004) of the links between the four cornerstones of health, the determinants of illness, and required responses to achieving wellbeing.

Figure A: Determinants and responses of the cornerstones of Te Whare Tapa Wha

Domain	Determinan	Response	Cornerston
Cultural	Colonisation: loss of autonomy	Political and cultural	Taha wairua
Emotional	Poor education, low self-	Management and support	Taha hinekaro
Social	Poverty: loss of land & food sources	Socio-economic development	Taha whanau
Physical	Lack of access to medical care	Access and treatment	Taha tinana

Determinants and responses of the cornerstones of Te Whare Tapa Wha: Adapted from Rochford, 1994.

Five Ways to Wellbeing

Five Ways to Wellbeing is a model that was developed by the New Economics Foundation (NEF) in 2008 in the United Kingdom. In developing the model, the aim was to identify the vital drivers of mental capital and wellbeing that are evidence-based, generic and appeal to a wide audience, target individuals rather than society, and varied so that actions remain fresh (New Economics Foundation, 2008). The five generic actions identified through a study undertaken are: **connect** with family friends, colleagues and neighbours; **be active** through physical activity commensurate with your level of fitness; **take notice** of what is going on around you wherever you are; **keep learning** about anything that captures your interest; and **give** whatever you have including time, smiles, kind words and of yourself to those around you (New Economics Foundation, 2008). The model proposes that when a person engages in these actions their wellbeing is enhanced resulting in a feeling of happiness that in turn has positive effect on mental capital (New Economics Foundation, 2008). According to the New Economics Foundation (2008), the actions are intended to have positive feedback in that, for instance, when you give to someone the gratitude expressed by that person will probably lead to the giver feeling satisfaction and a desire to continue giving.

References

Abma, T. A., & Widdershoven, G. A. M. (2008). Sharing stories: Narrative and dialgoue in responsive nursing evaluation. *Evaluation & the Health Professions, 28*(1), 90-109.

Ackerly, B., & True, J. (2008). Reflexivity in practice: Power and ethics in feminist research on international relations. *International Studies Review, 10*(4), 693-707.

Adult and Community Education Aotearoa. (2014a). *Maori success as Maori: A summary report from findings of the Maori success as Maori research project*. Wellington, New Zealand: Author.

Adult and Community Education Aotearoa. (2014b). *Pasifika success as Pasifika: Pasifika conceptualisation of literacy for successs as Pasifika in Aotearoa New Zealand*. Wellington, New Zealand: Author.

Advisory Committee on Official Statistics. (2009). *Good practice guidelines for the development and reporting of indicators*. Wellington, New Zealand: Statistics New Zealand. Retrieved from www.statisphere.govt.nz/further...and.../GoodPracticeGuidelines.pdf

Antrop, M. (2000). Background concepts for integrated landscape analysis. *Agriculture, Ecosystems and Environment, 77*, 17-28.

Atkinson, P., & Coffey, A. (2004). Texts: Analysing documentary realities. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice* (2nd ed.). London, England: Sage.

Aubel, J. (1999). *Participatory program evaluation manual: Involving program stakeholders in the evaluation process* (2nd ed.). Calverton, MD: USAID.

Aune, J. B. (2000). Logical Framework Approach and PRA - mutually exclusive or complementary tools for project planning? *Development in Practice, 10*(5), 687-690. doi: 10.1080/09614520020008850

Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, CA: Thomson Wadsworth.

Bailey, J. (2008). First steps in qualitative data analysis: Transcribing. *Family Practice Advance*, 127-131. doi: 10.1093/fampra/cmn003

Bakewell, O., & Garbutt, A. (2005). *The use and abuse of the Logical Framework Approach*. Oxford, England: INTRAC.

Baur, V. E., Abma, T. A., & Widdershoven, G. A. M. (2010). Participation of marginalised groups in evaluation: Mission impossible? *Evaluation and Program Planning, 33*(3), 238-245.

Berg, B. L. (2007). *Qualitative research methods for social sciences* (6th ed.). Boston, MA: Allyn and Bacon.

Bourdieu, P., & Wacquant, L. J. D. (1992). *An invitation to reflexive sociology*. Chicago, IL: University of Chicago Press.

Brewer, J. D. (2000). *Ethnography*. Buckingham, UK: Open University Press.

Brisolara, S. (1998). The history of participatory evaluation and current debates. *New Directions for Evaluation, 80*(Winter), 25-41.

Brunner, I., & Guzman, A. (1989). Participatory evaluation: A tool to assess projects and empower people. In R. F. Conner, & M. Hendricks (Eds.), *International innovations in evaluation methodology: New directions for program evaluation* (Vol. 42, pp. 9-19). San Francisco, CA: Jossey-Bass.

Burgess, R. (1984). *In the field: An introduction to field research*. London, UK: George Allen & Unwin.

Campilan, D. M. (2000). Conceptual tools for tracking change: Emerging issues and challenges. In M. Estrella, J. Blauert, D. Campilan, J. Gaventa, J. Gonsalves, I. Guijt, . . . R. Ricafort (Eds.), *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa, ON: International Development Research Centre.

Carruthers, G., & Tinning, G. (2003). Where, and how, do monitoring and sustainability indicators fit into environmental management systems? *Australian Journal of Experimental Agriculture, 43*, 307-323.

Chambers, R. (1994). The origins and practice of participatory rural appraisal. *World Development, 22*(7), 953-969.

Chambers, R. (1997a). *Whose Reality Counts? Puttin the First Last.* . London, UK: Intermediate Technology.

Chambers, R. (1997b). *Whose reality counts? Putting the first last*. London, UK: Intermediate Technology.

Chouinard, J., & Cousins, J. B. (2009). A review and synthesis of current research on cross-cultural evaluation. *American Journal of Evaluation, 30*, 457-494. doi: 10.1177/1098214009349865

Christchurch City Council. (2010). Shirley/Papanui 2008/2009 Strengthening Communities Fund, Small Grants Fund, and Discretionary Response Fund - End of project accountability reports. Retrieved 12 January 2015, from http://resources.ccc.govt.nz/files/thecouncil/meetingsminutes/agendas/2010/april/shirleyp apanui14th/clause12scfsgfdrsaccountability.pdf

Cleaver, F. (2001). Institutions, agency and the limitations of participatory approaches to development. In B. Cooke, & U. Kothari (Eds.), *Participation: The new tyranny?* London, UK: Zed Books.

Community Indicators Consortium. (n.d.). Indicator projects. Retrieved 15 December 2014, from http://www.communityindicators.net/projects

Community Waitakere. (2012a). *Case studies - Evaluation frameworks in community development organisations*. Auckland, New Zealand: Author.

Community Waitakere. (2012b). *Community development evaluation measures: Indicators of success*. Auckland, New Zealand: Author.

Community Waitakere. (2013a). *A community development assessment of strengths, aspriations, needs and opportunities of the community sector of Waitakere*. Auckland, New Zealand: Author.

Community Waitakere. (2013b). *Retrospective impact analysis: 7 case studies, looking back to move forward*. Auckland, New Zealand: Author.

Cooke, B., & Kothari, U. (2001). The case of participation as tyranny. In B. Cooke, & U. Kothari (Eds.), *Participation: The new tyranny?* London, UK: Zed Books.

Cousins, J. B., & Chouinard, J. A. (2012). *Participatory evaluation up close: An integration of research-based knowledge*. Charlotte, NC: Information Age Publishing.

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*(4), 397-418.

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation, 80*(Winter), 5-21.

Crawford, P., & Bryce, P. (2003). Project monitoring and evaluation: A method for enhancing the efficiency and effectiveness of aid project implementation. *International Journal of Project Management, 21*(5), 363-373.

Crishna, A. (2006). Participatory evalaution - (I) - Sharing lessons from fieldwork in Asia. *Child: Care, health and development, 33*(3), 217-223.

Cullen, A. E., Coryn, C. L. S., & Rugh, J. (2011). The politics and consequences of including stakeholders in international development evaluation. *American Journal of Evaluation, 32*(3), 345-361. doi: 10.1177/1098214010396076

Dart, J., & Davies, R. (2003). A dialogical, story-based evaluation tool: The Most Significant Change Technique". *American Journal of Evaluation, 24*(2), 137-155. doi: 11777/109821400302400202

DeGroff, A., Schooley, M., Chapel, T., & Poister, T. H. (2010). Challenges and strategies in applying performance measurement to federal public health programs. *Evaluation and Program Planning, 33*, 365-372.

Department for International Development. (2011). *Guidance on using the revised logical framework*. London, England: Author.

Department of Internal Affairs. (2011). *Review of selection New Zealand Government-funded community development programmes*. Wellington, New Zealand: Department of Internal Affairs.

Donabedian, A. (1998). The quality of care: How can it be assessed? *JAMA, 26D*(12), 1743-1748.

Donaldson, S., & Scriven, M. (2003). *Evaluating social programs and problems: Visions for the new millennium*. Mahwah, NJ: Lawrence Erlbaum Associates.

Doreen, W. V., Buckaert, G., & Halligan, J. (2010). *Performance management in the public sector*. Oxon, United Kingdom: Routledge Ltd.

Drucker, P. (1955). *The practice of management*. London, United Kingdom: First Collins Ltd.

Duignan, P., Casswell, S., Howden-Chapman, P., Barnes, H. M., Allan, B., & Conway, K. (2003). *Community project indicators framework: Its use in community projects*. Wellington, New Zealand: Ministry of Health.

Durie, M. (2004). *An indigenous model of health promotion*. Paper presented at the 18th World Conference on Health Promotion and Health Education, Melbourne, Australia.

Durie, M. (2011). Indigenizing mental health services: New Zealand experience. *Transcultural Psychiatry, 48*(1-2), 24-36.

Elasha, B. O., Elhassan, N. G., Ahmed, H., & Zakieldin, S. (2005). *Sustainable livelihood approach for assessing community resilience to climate change: case studies from Sudan*. Working Paper, (17). Assessments of Impacts and Adaptations to Climate Change.

Estrella, M., Blauert, J., Campilan, D., Gaventa, J., Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa, ON: International Development Research Centre.

Estrella, M., & Gaventa, J. (1998). *Who counts reality? Participatory monitoring and evaluation: A literature review*. Brighton, England: Institute of Development Studies, University of Sussex.

Evans, B., & Clark, A. (2012). *Evaluation report Seven Oaks Community Garden*. Christchurch, New Zealand: Project Lyttelton.

Fetterman, D. (2002). Empowerment evaluation: Building communities of practice and a culture of learning. *American Journal of Community Psychology, 30*(1), 89-102.

Flick, U. (2004). Constructivism. In U Flick, E. vonKardoff, & I. Steinke (Eds.), *A companion to qualitative research*. Thousand Oaks, CA: Sage.

Fraser, E. D. G., Dougill, A. J., Mabee, W. E., Reed, M., & McAlpine, P. (2006). Bottom up and top down: Analysis of participatory processes for sustainability indicator identification as a pathway to community empowerment and sustainable environmental management. *Journal of Environmental Management, 78*, 114-127.

Freebairn, D., & King, C. (2003). Reflections on collectively working toward sustainability: Indicators for indicators! *Australian Journal of Experimental Agriculture, 43*(223-238).

Friedman, M. (2005). *Trying hard is not good enough: How to produce measurable improvements for customers and communities*. Victoria BC, Canada: Trafford Publishing.

Gasper, D. (2000). Evaluating the 'Logical Framework Approach' towards learning-oriented development evaluation. *Public Admin. Dev., 20*, 17-28.

Gerring, J. (2004). What Is a Case Study and What Is It Good for? *The American Political Science Review, 98*(2), 341-354. doi: 10.2307/4145316

Giddens, A. (1991). *Modernity and self-identity*. Cambridge, UK: Polity.

Local Government Act (2002).

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues*. Thousand Oaks, CA: Sage.

Guijt, I. (2000). Methodological issues in participatory monitoring and evaluation. In M. Estrella, J. Blauert, D. Campilan, J. Gaventa, J. Gonsalves, I. Guijt, . . . R. Ricafort (Eds.), *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa, ON: International Development Research Centre.

Hamilton, C., Rai, R. K., Shrestha, R. B., Maharjan, M., Rasaily, L., & Hood, S. (2000). Exploring visions: Self-monitoring and evaluation processes within the Nepal-UK Community Forestry Project. In M. Estrella, J. Blauert, D. Campilan, J. Gaventa, J. Gonsalves, I. Guijt, . . . R. Ricafort (Eds.), *Learning from change: Issues and experiences in participatory monitoring and evaluation* (pp. 15-31). Ottawa, ON: International Development Research Centre.

Handley, S. (2008). Neighbourhood centres and results accountability: A conversation with Mark Friedman. *Local: The Newsletter for Community Development in NSW, Autumn*, 1-28.

Harmsworth, G., Young, R., Walker, D., Clapcott, J., & James, T. (2011). Linkages between cultural and scientific indicators of river and stream health. *New Zealand Journal of Marine and Freshwater Research, 45*(3), 423-436.

Hayes, R., & Oppenheim, R. (1997). Constructivism: Reality is what you make it. In T. Sexton, & B. Griffin (Eds.), *Constructivist thinking in counseling practice, research and training* (pp. 19-41). New York, NY: Teachers College Press.

Hillyer, K., & Purohit, S. (2007). Moving forwards with participatory monitoring and evaluation. *Participatory Learning and Action, 56*, 14-20.

Holte-McKenzie, M., Forde, S., & Theobald, S. (2006). Development of a participatory monitoring and evaluation strategy. *Evaluation and Program Planning, 29*, 365-376.

Hummelbrunner, R. (2010). Beyond logframe: Critique, variations and alternatives. In N. Fujita (Ed.), *Beyond logframe: Using systems concepts in evaluation*. Tokyo, Japan: Foundations for Advanced Studies on International Development.

Independent Māori Statutory Board. (n.d). *The Māori plan for Tāmaki Makaurau*. Auckland, New Zealand: Author.

Innes, J. E., & Booher, D. E. (1999). *Indicators for sustainable communities: A strategy building on complexity theory and distributed intelligence*. Working Paper 99-04. Institute of Urban and Reginal Development, University of California. Berkelely, CA.

Izurieta, A., Petheram, L., Stacey, N., & Garnett, S. T. (2013). Cost of participatory monitoring and evaluation of joint management of protected areas in the Northern Territory, Australia. *Australasian Journal of Environmental Management, 20*(1), 21-33.

Izurieta, A., Sithole, B., Stacey, N., Hunter-Xenie, H., Campbell, B., Donohoe, P., . . . Wilson, L. (2011). Developing indicators for monitoring and evaluating joint management effectivenesss in protected areas in the Northern Territory, Australia. *Ecology and Society, 16*(3). doi: http://dx.doi.org/10.5751/ES-04274-160309

Jacobs, A., Barnett, C., & Ponsford, R. (2010). Three approaches to monitoring: Feedback systems, participatory monitoring and evaluation and logical frameworks. *IDS Bulletin, 41*(6), 36-44.

Jacobson, S. K. (1991). Evaluation model for developing, implementing, and assessing conservation education programs: Examples from Belize and Costa Rica. *Environmental Management, 15*(2), 143-150.

Johnson, N., Lilja, N., Ashby, J. A., & Garcia, J. A. (2004). The practice of participatory research and gender analysis in natural resource management. *Natural Resources Forum, 28*, 189-200.

Johnson, P. (2004). Analytic induction. In C. Cassell, & G. Symon (Eds.), *Essential guide to qualitative methods in organizational research*. London, UK: Sage.

Keevers, L., Treleaven, L., Sykes, C., & Darcy, M. (2012). Made to measure: Taming practices with results-based accountability. *Organization Studies, 33*(1), 97-120.

Kibuga, K. F., Bibby, S., & Alfred, S. (1999). Participatory evaluation of a community water project in Tanzania. *PLA Notes, 1999*(35), 70-73.

King, C., Gunton, J., Freebairn, D., Coutts, J., & Webb, I. (2000). The sustainability indicators industry: Where to from here? A focus group study to explore the potential of farmer participation in the development of indicators. *Australian Journal of Experimental Agriculture, 40*, 631-642.

King, J. A., & Ehlert, J. C. (2008). What we learned from three evaluations that involved stakeholders. *Studies in Educational Evaluation, 34*(4), 194-200.

King, N. (2004). Using templates in the thematic analysis of text. In C. Cassell, & G. Symon (Eds.), *Essential guide to qualitative methods in organizational research* (pp. 256-270). London, UK: Sage.

Kvale, S. (2006). Dominance through interviews and dialogues. *Qualitative Inquiry, 12*, 480-500. doi: 10.1177/1077800406286235

Lather, P. (1991). *Getting smart: Feminist research and pedagogy within the postmodern*. New York, NY: Routledge.

Lennie, J., Tacchi, J., Koirala, B., Wilmore, M., & Skuse, A. (2011). Equal access particpatory monitoring and evaluation toolkit: Helping communication for development organisations to demonstrate impact, listen and learn, and imprve their practices. Retrieved 14 January, 2015, from

http://betterevaluation.org/sites/default/files/EA_PM%26E_toolkit_front_pages%26introduction_for_publication.pdf

Leonard, L., & Memon, A. (2008). *Community outcome processes as a forum for community governance*. Hamilton, New Zealand: The International Global Change Institute, University of Waikato.

Levermore, R. (2011). Evaluating sport-for-development: Approaches and critical issues. *Progress in Development Studies, 11*(4), 339-353.

Lindenmayer, D. B., Gibbons, P., Bourke, M., Burgman, M., Dickman, C. R., Ferrier, S., . . . Zerger, A. (2012). Improving biodiversity monitoring. *Austral Ecology, 37*, 285-294. doi: 10.1111/j.1442-9993.2011.02314.x

Loomis, T. M. (2011). *Community-led development in Wairoa: An account of the 2009-10 Wairoa Social Development Project*. Wairoa, New Zealand: Wairoa Community Deveopment Trust.

Lunt, N., Davidson, C., & McKegg, K. (2003). *Evaluating policy and practice: A New Zealand reader*. Auckland, New Zealand: Pearson Education.

Madill, A., Jordan, A., & Shirley, C. (2000). Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British Journal of Psychological, 91*, 1-20.

Make it Happen Te Hiku. (2014). *Make it happen Te Hiku: Community action plans for outcomes based social development*. Wellington, New Zealand: Ministry of Social Development.

Marr, B. (2009). *Managing and delivering performance*. Amsterdam, The Netherlands: Butterworth-Heinemann.

Marsden, D., & Oakley, P. (1990). *Evaluating social development projects*. OXFAM. London, UK.

Mathison, S. (2005). Outcomes. *Encyclopedia of Evaluation* (pp. 288). Thousand Oaks, CA: Sage Publications, Inc.

McDuff, M. (2001). Building the capacity of grassroots conversation organizations to conduct participatory evaluation. *Environmental Management, 27*(5), 715-727.

McNay, L. (2000). *Gender and agency: Reconfiguring the subject in feminist and social theory*. Cambridge, UK: Polity.

Memon, A., & Johnston, K. (2008). Institutional barriers to developing community indicators in New Zealand: A preliminary assessment. *Commonwealth Journal of Local Governance, 1*, 70-91.

Ministry of Social Development. (n.d.). Results based accountability. Retrieved 12 January 2015, from http://www.familyservices.govt.nz/working-with-us/funding-and-contracting/results-based-accountability/

Nelson, K. D., & Tipa, G. T. (2012). *Cultural indicators, monitoring frameworks and assessment tools*. Wheel of Water Research Programme. Retrieved from http://comar.co.nz/wp-content/uploads/Report-for-Wheel-of-Water-Project-cultural-tools-in-use-and-development.pdf

New Economics Foundation. (2008). *Five ways to wellbeing*. London, United Kingdom: Author. Retrieved from http://b.3cdn.net/nefoundation/8984c5089d5c2285ee_t4m6bhqq5.pdf

Njuki, J., Mapila, M., Kaaria, S., & Magombo, T. (2008). Using community indicators for evaluating research and development projects. *Development in Practice, 18*(4/5), 633-642.

O'Leary, Z. (2010). *The essential guide to doing your research project*. Los Angeles, CA: Sage.

Organisation X. (2012). *Strategic plan: 2012-2015*. Author.

Parkinson, S. (2009). Power and perceptions in participatory monitoring and evaluation. *Evaluation and Program Planning, 32*, 229-237.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Los Angeles, CA: Sage.

Pretty, J. N. (1995). *Regenerating agriculture: Policies and practices for sustainability and self-reliance*. Washington, DC: Joseph Henry Press.

Pretty, J. N., & Vodouhê, S. D. (1998). Using rapid or participatory rural appraisal. In B. Swanson, R. Bentz, & A. Sofranko (Eds.), Improving agricultural extension: A reference manual. Rome, Italy: Food and Agricultural Organisation. Retrieved from http://www.fao.org/docrep/W5830E/W5830E00.htm.

Prinsen, G. (2013). *Measuring the impact of a kiss on love: Designing indicators for multi-dimensional concepts such as sovreignty, resilience, sustainability, or fragility*. Paper presented at the Sustainable Development in Oceania - Towards a New Ethic?, Kone, New Caledonia.

Reed, M. S., & Dougill, A. J. (2002). Participatory selection process for indicators of rangeland condition in the Kalahari. *The Geographical Journal, 168*(3), 224-234.

Reed, M. S., Fraser, E. D. G., & Dougill, A. J. (2006). An adaptive learning process for developing and applying sustainability indicators with local communities. *Ecological Economics, 56*, 406-418.

Rist, R. (1995). Postscript: Development questions and evaluation answers. *New Directions for Program Evaluation, 67*, 167-174.

Ritchie, M. (2000). *Decision support systems for managing rangelands.* Paper presented at the Changing landscapes - shaping futures. Proceedings of the international landcare 2000 conference, Melbourne Convention Centre, Melbourne.

Roche, C. (1999). *Impact assessment for development agencies: Learning to value change*. Oxford, UK: Oxfam GB.

Rochford, T. (2004). Whare Tapa Whara: A Māori model of unified health. *The Journal for Primary Prevention, 25*(1), 41-57.

Save the Children UK. (2008). *Menu of outcome indicators*. London, UK: Author.

Scheyvens, R. (2007). The idiosyncrasies of indicators. *DevForum, 27*, 9-12.

Schischka, J. (2009). *Te Whare Roimata - Participatory evaluation of a New Zealand urban indigenous development agency*. Paper presented at the Human Development and Capability Association Annual Conference, Lima, Peru.

Simons, H. (2009). *Case study research in practice*. London, England: : Sage.

Somers, C. (2005). Evaluation of the Wonders in Nature-Wonders in Neighborhoods conservation education program: Stakeholders gone wild! *New Directions for Evaluation, 108*, 29-46.

Springer-Heinze, A., Hartwich, F., Henderson, J. S., Horton, D., & Minde, I. (2003). Impact Pathway Analysis: An approach to strengthening the impact orientation of agricultural research. *Agricultural Systems, 78*, 267-285.

Stacey, N., Izurieta, A., & Garnett, S. T. (2013). Collaborative measurement of performance of jointly managed protected areas in northern Australia. *Ecology and Society, 18*(1). doi: http://dx.doi.org/10.5751/ES-05273-180119

Statistics New Zealand. (n.d). *Indicator guidelines*. Author. Retrieved from http://www.stats.govt.nz/methods/indicator-guidelines.aspx

Stoneham, G., Eigenraam, M., Ridley, A., & Barr, N. (2003). The application of sustainability concepts to Australian agriculture: An overview. *Australian Journal of Experimental Agriculture, 43*, 195-203.

Taylor, S. J., & Bogdan, R. (1984). *Introduction to qualitative research methods: The search for meanings* (2nd ed.). New York, NY: Wiley.

Terry, A. (2008). Community sustainable-development indicators: A useful participatory technique or another dead end? *Development in Practice, 18*(2), 223-234.

Thayer, C. E., & Fine, A. H. (2001). Evaluation and outcome measurement in the non-profit sector: Stakeholder participation. *Evaluation and Program Planning, 24*(1), 103-108.

The Scottish Government. (n.d.). Using the LEAP Framework: Developing outcome indicators. Retrieved 06 October, 2014, from http://www.scdc.org.uk/media/resources/what-we-do/leap/Developing%20indicators.pdf

Trotman, R. (2008). *Promoting goodness: A guide to evaluating programmes and projects*. Te Pou o Te Whakaaro. Retrieved from http://www.tepou.co.nz/assets/images/content/training_funding/tools-for-learning/files/Promoting%20Goodness.pdf

Tuhiwai-Smith, L. (1999). *Decolonising methodologies*. London: England: Zed Books.

UNDP. (1990). *Human development report*. New York, NY: Author.

UNDP. (2002). *Handbook on monitoring and evaluating for results*. New York, NY: Author.

UNDP. (2009). *Handbook on planning, monitoring and evaluating for development results*. New York, NY: Author.

UNDP. (n.d). *Selecting indicators for impact evaluation*. Retrieved from http://www.ngoconnect.net/documents/592341/749044/Selecting+Indicators+for+Impact+Evaluation

United Nations. (n.d.). About MDGs. Retrieved 15 December 2014, from http://www.unmillenniumproject.org/goals/

Vernooy, R., Qiu, S., & Jianchu, X. (2006). The power of participatory monitoring and evaluation: Insights from south-west China. *Development in Practice, 16*(5), 400-411.

Wallace, T. (2010). *Monitoring and Evaluation in development work: Why involving local people matters and how it can be done*. Paper presented at the The 6th International Colloquium of Working Equids: learning from others: Proceedings on an International Colloquium, New Delhi, India.

Warsame, M., Mortensen, A., & Janif, J. (2014). Improving educational achievement for students from Somali backgrounds in Auckland, New Zealand. *Australasian Review of African Studies, 35*(2), 25-44.

Weaver, L., & Cousins, J. B. (2001). Unpacking the participatory process. *Journal of Multidisciplinary Evaluation, 1*(1), 19-40.

Willets, J., & Crawford, P. (2007). The most significant lessons about the Most Significant Change technique. *Development in Practice, 17*(3), 367-379.

Woodcock, A., Cronin, O., & Forde, S. (2012). Quantitative evidence for the benefits of Moving the Goalposts, a sport for development project in rural Kenya. *Evaluation and Program Planning, 35*, 370-381.

World Bank. (n.d.). Indicators. Retrieved 15 December 2014, from
http://data.worldbank.org/indicator

World Meteorological Organization. (2011). *Guidelines for implementing a quality management system in hydrology*. Retrieved from
http://www.wmo.int/pages/prog/hwrp/chy/chy13/documents/awg3/Doc6_2-Quality_Management.pdf

Yin, R. K. (1984). *Case study research: Design and methods*. Beverley Hills, CA: Sage.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Los Angeles, CA: Sage.