University of Strathclyde

Department of Management Science

Putting (programme) theory into practice: the use of programme theory in programme evaluation practice in small third sector organisations

Current practice and future improvements

Robyn Millar

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

2020

Dedication

To Papa George,

To Papa Will (who passed away in July 2019),

For two loving Papas, who taught me what it meant to work hard, to always try your best, and to be grateful for those who support you along the way.

Declaration of authenticity and author's rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date:

Acknowledgements

There are several people I wish to acknowledge, who have supported me throughout the PhD process.

Thanks must first go to my two supervisors, Prof Alec Morton & Prof Sir Harry Burns for their support and encouragement. In particular, I would like to express my thanks to Prof Alec Morton, for encouraging me to independently define a research project that interests and motivates me, and for continually providing constructive and timely feedback on my work. I would also like to express my thanks to Alec for offering me many opportunities to be involved in university life, outside of PhD study, through various administrative duties, research projects, and knowledge exchange activities. My PhD experience certainly would not have been the same without these opportunities.

Particular thanks must go to all those who participated in this research. Without you, I could not have learned as much as I did.

This work was supported by the Economic and Social Research Council [award reference 1684283]. ESRC not only provided a funding source for my PhD studies, but also access to a great community of other PGR students and learning opportunities through SGSSS.

Thanks must also go to my family, in particular, Mum, Dad, Molly, and Lucy. You have provided continual encouragement and support in everything I do, not just over the last four years.

To Conor, who has had to live with me (and listen to me) for the last four years. Throughout this process, you have always listened to my worries and kept me grounded, whilst also making sure I enjoyed the things going on outside of PhD life.

To Mary, thank-you for meticulously reading my work, and offering constructive guidance on the use of a comma. I am forever grateful for the time you have devoted to helping me edit such a large piece of work.

To all my friends, thank-you. I would especially like to thank all those who I met through ‘PhD Women in Scotland’ and writing retreat events. I could never have imagined how much these social networks would mean to me during this process.

Abstract

The theory-practice gaps in the field of evaluation demand evaluation. Research on evaluation aims to address these gaps. Specifically, this thesis focuses on the use of programme theory in the context of programme evaluation practice in small third sector organisations. In this context, social programming efforts and programme evaluation are critical to ensure the effective design, and implementation of programmes and services for the individuals, groups, and communities that small third sector organisations work with. Despite the importance of social programming and programme evaluation in this setting, there are numerous contextual challenges facing programme evaluation practice.

This thesis highlights that how we conceptualise evaluation practice has implications for how we might go about conducting research on evaluation practice: such implications relate to the focus of the research and the methodological framework used. I find that the literature on the use of programme theory in evaluation practice is not consistent with our understanding of programme evaluation practice in small third sector organisations. As such, I use the empirical component of this thesis to explore the use of programme theory in the context of programme evaluation practice in small third sector organisations from a 'practitioner-oriented perspective'. I address two research questions about current practice in the use of programme theory in small third sector organisations, and about how the development of programme theory can improve future programme evaluation practice in this setting. I adopt a multi-methodological framework to address the research questions.

In terms of current practice, I find that much of the understanding and use of programme theory in small third sector organisations can be considered as tacit in nature. I also find that, whilst there may be challenges in making knowledge which is tacit in nature, explicit, that programme theory is not used explicitly in formal, or systematic, programme evaluation activities. This is also partly due to other factors including, practical constraints, and the fact that the

use of programme theory is not explicitly asked for by funders. In terms of improving future practice, I find that the development of programme theory has two primary roles in helping evaluators think more purposefully and constructively about programme implementation and outcome evaluation: that is, in its role as a 'tool for evaluative thinking', through its confirmatory function and visual nature.

Overall, this thesis contributes to our understanding of the use of programme theory in small third sector organisations in terms of the better conceptualising current practice, as well as understanding how the development of programme theory can improve future programme evaluation practices in this setting. These contributions have implications for programme evaluation theory and practice in small third sector organisations, and for research on evaluation more generally.

Table of Contents

LIST OF FIGURES, TABLES, AND IMAGES

Figures

Tables

Images

CHAPTER 1 INTRODUCTION

1.1 Introduction

This thesis explores the use of programme theory in programme evaluation practice in small third sector organisations. The focus of this research on programme evaluation practice means that it falls within the field of 'research on evaluation'. The purpose of this introductory chapter is to provide context for, and understanding of, the research area, as well as the context in which the empirical component of this thesis is based. Overall, this chapter aims to highlight the significance and relevance of the empirical component of the thesis.

Specifically, in this chapter, I describe my personal and the academic rationale for conducting this research. The academic rationale first highlights the importance of programme evaluation, as well as research on programme evaluation practice. Second, the academic rationale also describes the context in which I focus the research, programme evaluation practice in small third sector organisations. Third, I provide a brief overview of the research problem identified (presented in more detail in Chapter 2) and the research approach taken to address the two key research questions (presented in more detail in Chapter 3). I briefly describe the methodology employed before outlining the contributions and structure of the thesis at the end of the chapter.

1.2 Personal rationale for the research

I became interested in the topic of programme evaluation practice whilst completing my Masters and during the first year of my PhD studies. The focus of my Master's thesis (and most of the first year of my PhD) was on the economic consequences of youth disadvantage with a view to exploring cost-effective strategies, or programmes, to tackle youth disadvantage. The evaluation of such strategies, or programmes, was of interest in terms of improving our understanding of the most cost-effective ways of tackling such disadvantage. I became particularly interested in understanding *how* and *why*

programmes worked to tackle social issues, rather than simply understanding whether outcomes changed. So, I then spent time understanding the principles of 'programme theory'[1], which was to become the focus of this thesis - but not in the way I originally anticipated.

During my Masters and first year of my PhD, I also had the opportunity to learn about, and discuss, some of the challenges facing organisations that implement programmes and services to tackle social issues such as youth disadvantage: these organisations are, in many cases, third sector organisations. One of the practical challenges facing these organisations is in evaluating the programmes and services they deliver and using the findings emerging from programme evaluation in constructive ways. Whilst some organisations had the capacity to seek academic support for doing programme evaluation, the opportunity for such organisations to seek such support for programme evaluation was limited. As a result, it was clear that much programme evaluation activity was conducted internally within these organisations by staff members who serve multiple roles within the organisation, e.g. programme management, service delivery, and administration. Simultaneously, I also noticed that these contextual challenges are not well reflected in the literature on the use of programme theory in programme evaluation practice.

As a result, I sought to continue with research which could support the practice of evaluation *within* organisations rather than conduct evaluation *on how* the programmes are delivered by the organisations. I considered that this aim could help improve the ability of these organisations to do programme

[1] Programme theory explicitly reflects the assumptions about how and why programmes work to change outcomes for programme participants, usually depicted in the form of a word-and-arrow diagram.

evaluation which would in turn lead to the delivery of more effective programmes and services for the beneficiaries[2] with whom they work.

1.3 Academic rationale for the research

In what follows, I describe the academic rationale for this research to provide understanding of the academic and practical relevance of this study. I first define programme evaluation, before commenting on the need for research on programme evaluation practice. I then describe the empirical context in which this research focuses, and the substantive focus on the use of programme theory in programme evaluation practice.

1.3.1 Programme evaluation and programme evaluation practice

As a broad field, 'evaluation' can refer to the evaluation of programmes, policies, products, personnel, performance, and proposals, or the 'Big Six' P's (Scriven, 1991). The 'Big Six' P's illustrate that the field of evaluation is trans-disciplinary in the sense that almost anything can be evaluated (Shadish, Cook and Leviton, 1991). For the purposes of this thesis, I focus on the evaluation of social programmes: social programmes have the shared goal of contributing to social betterment[3] (Mark, Henry and Julnes, 2000), i.e. improving some social outcome. At a general level, *programme* evaluation includes the activities by which one can make sense of programmes through conducting systematic inquiry to describe or explain the operations, outputs, outcomes, and social implications of programmes[4] (Mark, Henry and Julnes, 1999, 2000).

---

[2] Beneficiaries refer to those individuals or groups who are intended to benefit from programmes and services offered by organisations such as small TSOs.

[3] Social betterment refers to the goal of reducing/preventing social problems, the improvement of social conditions, and in general the alleviation of human suffering (Mark, Henry and Julnes, 2000, p.7).

[4] Programme evaluation activities can be formative (carried out at the beginning of or during a programme) or summative (carried out at the end of a programme). These activities can focus on implementation of the programmes, the outcomes of the programmes, the resource to

Such activities can be conducted at any stage in the programme life-cycle using a range of methods and approaches (Mark, Henry and Julnes, 1999, 2000).

To take an example, one could evaluate the implementation of a programme, i.e. how the activities or services were delivered. To do this, an evaluation might seek to observe programme activities in action. Alternatively, or in addition, an evaluation could interview programme beneficiaries to understand their perspectives on how a programme was delivered, and then compare this with how the programme was intended to be delivered. The evaluation of programme implementation could be undertaken during the course of the programme, i.e. formatively, to help understand how things could be improved. Alternatively, it could be undertaken after a programme had been delivered, i.e. summatively, to assess the extent to which the programme was implemented as intended, otherwise referred to as assessing programme fidelity.

Programme evaluation can also serve several purposes within organisations delivering programmes. It is commonly considered to serve four key purposes: assessment of merit or worth (or judging the value of a programme), accountability (or oversight and compliance), knowledge development (or generating or testing theories and hypotheses in the context of programmes), and programme/organisational improvement (to change or develop programmes) (Mark, Henry and Julnes, 2000; Chelimsky, 2006). Overall, programme evaluation therefore plays a critical role in organisations in helping to achieve social betterment through their ability to “better select, oversee, improve, and make sense of programmes” (Mark, Henry and Julnes, 2000, p.3), or through learning about different programme strategies more generally.

benefit ratio of the programme, or the overall impact of a programme (Langbein, 2012; Center of Disease Control and Prevention, 2019)

In terms of defining research on evaluation, programme evaluation practice is often cited as the 'doing of' programme evaluation, or how evaluators conduct their work (Rog, 2015). However, this conceptualisation of evaluation practice is not helpful in terms of defining purposeful and systematic empirical inquiries on evaluation practice for a variety of reasons. First, it does not address *who* is practising evaluation i.e. who is conducting programme evaluation activities. Second, it does not indicate *what* exactly is being practised, i.e. the actual evaluation activities conducted. Third, it does not conceptualise *how* evaluation is being practised. Last, it does not address the contextual influences on programme evaluation practice, i.e. *why* programme evaluation practice is conducted the way that it is. I argue that considering these questions regarding the who, what, how, and why of evaluation practice provides a more refined conceptualisation of evaluation practice. Using a more refined conceptualisation of evaluation practice can not only help to identify gaps in our knowledge about evaluation theory, but it can also influence how we might go about conducting research on evaluation practice. I explain this, in more detail, in the following sections.

1.3.2 Research on programme evaluation practice

Broadly defined, research on evaluation is, "any purposeful, systematic, empirical inquiry intended to test existing knowledge, contribute to existing knowledge, or generate new knowledge related to some aspect of evaluation processes or products, or theories, methods, or practices" (Coryn et al. 2016). In this research, I use the term evaluation theory, as an umbrella term, to refer to evaluation processes, products, theories, methods, or approaches – I have defined this in the shaded box below. In the case of *programme* evaluation, research on evaluation can focus on the processes or products of *programme* evaluation and the theories, methods, or approaches that are intended to guide or inform *programme* evaluation practice, rather than the evaluation of any of

the other “Big 6” Ps[5]. From this point, I use the term ‘evaluation theory’ as defined below.

DEFINITION

Evaluation theory relates to the processes, products, theory, methods, or approaches that underpin the field of evaluation. Evaluation theory can be descriptive, i.e. it can describe evaluation practice, or prescriptive, i.e. it lays out how evaluation *should* be carried out. Exploring the gaps between descriptive and prescriptive theory can help to develop evaluation theory such that it supports the conduct of evaluation that is feasible and useful for evaluation practitioners, and as such, can inform a more effective dialogue between theory and practice in the evaluation field.

Importantly, this definition of research on evaluation defines research as *purposeful, systematic,* and *empirical* which distinguishes research on evaluation from single case reports of empirical evaluations which dominate the programme evaluation literature (Galport and Galport, 2015). Single cases of evaluations are neither systematic nor purposeful because, in most cases, their central focus is not the evaluation process, product, theory, method or practice. Rather, the purpose of single cases of evaluation is to report upon that specific evaluation and its findings as opposed to making purposeful comments about the approach taken, or the theory employed. Variable styles of reporting upon single case studies of evaluations can therefore make it difficult to draw out learning about evaluation theory from across studies.

The overall purpose of conducting research on evaluation is to close the theory-practice gap in the field of programme evaluation. It is a well-accepted notion that evaluation practice is not simply evaluation theory enacted (Dahler-

[5] E.g. policies, products, personnel, performance, and proposals (Scriven, 1991).

Larsen, 2014) because in practice, evaluation situations are rarely perfectly suitable for following a particular theory to the letter (Smith, 1993) due, in part, to the variety of contextual influences on evaluation practice for example. This notion reflects wider criticisms on management research more generally (Mintzberg, 1973): in such criticisms, scholars argue that sufficient research attention needs to be given to the task of managing and implementing managerial solutions, rather than simply focusing on developing such solutions. The same goes for evaluation practice. In particular, research on evaluation seeks to identify the gaps between descriptive and prescriptive theory in programme evaluation. This means that research on evaluation focuses on both descriptive theory – theory that describes how evaluation is conducted – and prescriptive theory – theories that explain how evaluation should be conducted (2003a). However, some argue that a lack of empirical studies on how evaluation is conducted, i.e. descriptive theory, means that the evaluation field is "built on, and constrained by prescriptive ideas" (Christie, 2003a, p.2). It is therefore imperative that research on evaluation attends to both descriptive and prescriptive components of evaluation theory to create a mutually informing dialogue between the two.

It follows that descriptive evaluation theory is central to the advancement of prescriptive theory in the field of programme evaluation (Smith, 1993). Empirical studies of descriptive evaluation theory help to understand under which conditions and contexts a prescriptive theory is suitable and functional and what can be achieved under those conditions (Cousins & Earl, cited in: Christie, 2003a, p.1). Research on evaluation practice should therefore endeavour to inform a dialogue between prescriptive and descriptive theory. In some cases, research on evaluation has effectively informed a dialogue between prescriptive and descriptive theory through the use of purposeful, systematic, and empirical inquiries. An example of this is the theory on 'evaluation use' where research on evaluation has informed understanding about how and why both the process of being involved in evaluation, and the findings of an evaluation, are useful to evaluation stakeholders (Johnson et al.,

2009). This research has helped to develop several popular frameworks for programme evaluation including Patton's *Utilisation-focussed evaluation*[6] (1997, 2008). This exemplifies the potential value and importance of research on evaluation.

However, despite calls for more research on evaluation, empirical studies have been lacking. For example, in 1993 Smith emphasised the need for more research on evaluation practice to better understand practice, to facilitate comparison of the feasibility and effectiveness of different evaluation models, methods and theories, and in order to understand the utility of prescriptive theory in actual practice (Smith, 1993). Despite this, more recent reviews of research on evaluation make similar calls to the academic evaluation community (Vallin, Philippoff, Pierce and Brandon, 2015; Coryn et al., 2017). The repetitive nature of these calls for more research on evaluation highlights that there must be some limiting factors for conducting such research.

There are several reasons why research on evaluation remains limited. An important reason is that there is often a lack of conceptual consensus within the field of evaluation, i.e. various meanings are commonly attached to key terms (King, 2003). For example, this thesis focuses on the use of programme theory in programme evaluation, yet it has been noted that there are multiple interlinked and overlapping terms to describe what is essentially the same thing as programme theory, or some slight variation on programme theory. Funnell and Rogers list over 20 terms which are the same as, or are at least closely related to, the concept of programme theory (Funnell and Rogers,

[6] Utilisation-focussed evaluation (UFE) is an approach to conducting evaluation that places the usefulness of the evaluation to evaluation stakeholders as the top priority for evaluation efforts (Patton, 1997, 2008)

2011)[7]. As such, it is easy to see how conceptual confusion could complicate the conduct of research on evaluation through: identifying which concepts research should focus on; how to locate examples of concepts if multiple terminologies exist; and how to address evaluation concepts when those using them may not even be aware of them or use different terms by which to talk about them.

Other reasons for the lack of research on evaluation also relate to practical issues in the field of evaluation. The pragmatic and professional, rather than academic, focus of evaluators, and the field of evaluation more broadly, mean that theory development remains a low priority (Christie, 2003a). There is also a relative lack of financial research support for research on evaluation because much evaluation is conducted in non-evaluation-specific academic fields (King, 2003) e.g. health, education, and social work. As such, there is less incentive to publish research on evaluation, in fields outside the evaluation field, where other priorities, such as generating programme-related theory, take priority. More generally, evaluation as an academic field is relatively young and as such may have consequences for the ability to develop and refine theory about it (King, 2003). As a result, there is a lack of frameworks to guide empirical research on evaluation. Therefore, the quality of research on evaluation is often weakened as it can be difficult to summarise across studies, especially those studies which are narrative or anecdotal in nature or those which are single case reports of empirical evaluations (Vallin et al., 2015).

Overall, despite the importance of conducting programme evaluation, and the many purposes it serves in terms of helping organisations to better select,

[7] For example, terms relating to programme theory according to Funnell and Rogers include but are not limited to: chains of reasoning, causal chain, causal map, impact pathway, intervention framework, intervention logic, intervention theory, logic model, logical framework, mental model, outcomes hierarchy, outcomes line, performance framework, programme logic, programme theory, reasoning map, results chain, theory of action, theory of change

oversee, improve, and make sense of programmes, there are calls for more systematic and purposeful empirical research on programme evaluation practice so as to better inform a dialogue between what evaluation theory says we should do and what we do, or can do as evaluators, within the contexts in which programme evaluation practice actually takes place. That being said, I now introduce the empirical context in which this research is set: programme evaluation practice in small third sector organisations.

1.3.3 Empirical context: small third sector organisations

The empirical context I am interested in is evaluation practice in small third sector organisations (herein 'small TSOs'). Whilst there is a range on views on what defines the 'third sector' (Alcock, 2010), I take TSOs to be those organisations outside of the predominant control of government, that serve a public and/or social purpose, and that pursue activities without compulsion (Salamon and Sokolowski, 2016). This can include, but is not limited to, charitable organisations, non-profit organisations, voluntary organisations, and social enterprises. In the context of this research, the focus on small TSOs corresponds to the fact that programme evaluation in smaller organisations is usually conducted by someone without formal training in programme evaluation and who conducts programme evaluation as but one part of their job. Naturally this corresponds to charities with lower incomes and smaller numbers of employees. There are several reasons why small TSOs are an interesting empirical context in which to consider research on programme evaluation practice. In what follows I outline why I think this is so in terms of both the importance of TSOs in wider society, as well as the programme evaluation practice context of small TSOs more specifically.

1.3.3.1 Importance of third sector in society

In the first instance, the importance of the third sector, particularly small TSOs, in helping to address some of the bigger challenges facing societies has been noted:

> *"If you want to do something that is very human centred, that has a connection to geography, if you want to tackle crime … you need people who come from that estate … if you want a human solution, you need a human-scaled charitable response and sadly, too often what we're seeing is the money moving into the generic, the larger charity, and away from the human, passionate, and specialist charity"*
>
> Duncan Shrubsole (2018)[8]

In many countries, the third sector is playing an increasingly important role in tackling the social problems faced by communities, e.g. social disadvantage, public health, and education. In the UK for example, third sector organisations, including social enterprises, are important stakeholders in the provision of programmes and services, and in generating social change, at both the local and national level, as well as in both urban and rural communities (Steiner and Teasdale, 2018). The important role played by these organisations in the UK can be seen in practice by the increasing contracting out of public services (Hardwick, Anderson and Cooper, 2015), and in academia by the increased research interest in the potential of TSOs to improve health and social outcomes (Roy, Donaldson, Baker and Kerr, 2014; Markantoni, Steiner and Meador, 2019). The increasingly central role played by TSOs is often cited as a consequence of their specialist knowledge of the needs of those individuals and groups within, and of communities themselves (Macmillan, 2010; Lang and Hardwick, 2016) and thus their enhanced ability to design programmes and services to meet local needs (Hardwick, 2018). This specialist knowledge is cited to be a consequence of the 'know-how', or tacit knowledge, of those working in TSOs. Such knowledge is personalised and is developed within, and integrated into, programme or services beneficiaries. This knowledge is found to be used to contribute to better programmes and services for beneficiaries, but is not necessarily made explicit for the purposes of programme evaluation (Hardwick, 2018).

[8] The Charity Business podcast (BBC, 2018) - https://www.bbc.co.uk/programmes/b09w2tjd

1.3.3.2 Programme evaluation practice in small TSOs

Alongside the increasingly important role played by TSOs in many contexts, there is an acknowledgement of, and interest in, the role of programme evaluation in such organisations (Fine, Thayer and Coghlan, 2000; Thayer and Fine, 2001; Carman and Fredericks, 2008, 2010; Arvidson, 2009; Ellis and Gregory, 2009; Carman, 2011; Harlock, 2013; Ricciuti and Calo, 2016; Bach-Mortensen and Montgomery, 2018). These studies tend to focus descriptively on overall programme evaluation practice and the factors perceived to facilitate, or act as barriers to, conducting useful evaluations (Arvidson, 2009). Studies also address the perceived roles that programme evaluation plays for TSOs e.g. as a resource drain, for external promotions, and for strategic planning (Carman and Fredericks, 2008). Other studies provide analyses of why third sector organisations conduct evaluation (Carman, 2011). Overall, studies tend to be descriptive in nature with a focus on programme evaluation at broad level, and often discuss programme evaluation in terms of reporting and accountability requirements.

In the context of this thesis, *small* TSOs present an interesting context of programme evaluation practice for a number of reasons which I outline below:

1. The internal and external organisational contexts of small TSOs organisations tackling social problems can present a challenge for those practising evaluation.

    a. In terms of the internal environment, TSOs are mostly small in size with respect to income and staffing, with nearly 97% of charities in the UK having an income less than £1 million per year[9]. Such financial and staffing constraints mean that programme evaluation

[9] https://www.ncvo.org.uk/about-us/media-centre/briefings/1721-fast-facts-about-the-charity-sector

is often conducted in-house rather than by hiring external evaluators. Moreover, this means that those carrying out evaluation in such organisations often have multiple roles e.g. programme development, management, service delivery, administration. The capacity of those practising evaluation in third sector organisations is therefore impacted in two ways. First, those practising evaluation are not necessarily formally trained in evaluation skills and as such perhaps do not know how to approach the task of evaluating. Second, given the small size and income of many such organisations, the capacity to practise evaluation is influenced in terms of operational resources such as finance and time (Ellis and Gregory, 2009).

b. In terms of the external environment, TSOs are also faced with increasingly complex funding and accountability environments (Ellis and Gregory, 2009) often characterised through principal-agent relations[10] whereby third sector organisations are required to meet funder expectations regarding performance. (Tenbensel, Dwyer and Lavoie, 2014). Moreover, third sector organisations need to demonstrate various types of accountability including providing information on operations, legitimacy, and social value (Costa, Ramus and Andreaus, 2011) in their relationships with funders, the state, communities and practitioners. However, empirical research shows that the capacity of TSOs to demonstrate multiple levels of accountability is limited (Costa, Ramus and Andreaus, 2011) and as a result, they are often forced to focus on economic and financial reporting (Costa, Ramus and Andreaus, 2011) to protect their own

[10] Principle-agent theory – agency contracts are those that presume that the agent is subordinate to the principal and that the principal is inherently sceptical of the agent's performance and as a result, the agent needs to show that its performance meets expectations (Tenbensel, Dwyer and Lavoie, 2014)

interests (Dhanani and Connolly, 2012) rather than focussing on social value. These increasingly complex, and often stringent, accountability requirements can dilute the mission and values of TSOs (Tenbensel, Dwyer and Lavoie, 2014). The missions and values of small TSOs are specifically that which characterise the sector as a whole.

Moreover, accountability requirements often lead to a perception of an imbalance in power, particularly in contracting relationships which can impact evaluation practice in third sector organisations (Ebrahim, 2003; Tenbensel, Dwyer and Lavoie, 2014). For example, performance monitoring contracts are often required from public sector contractors which somewhat contradict the reason why the third sector organisation was contracted to provide the service in the first place, i.e. that they were better placed to provide a service, over a public sector provider for example. An increase in accountability requirements can also decrease the organisation's ability to be accountable to other important stakeholders e.g. programme beneficiaries (Tenbensel, Dwyer and Lavoie, 2014). Overall, this imbalance in power can result in 'punishment' of third sector organisations through cutting funding, imposing conditions, and tarnishing reputations (Ebrahim, 2003).

2. The actual social issues that third sector organisations aim to tackle present challenges in terms of the practice of evaluation. For example, the multiplicity of the social issues being targeted is a challenge. Social problems are often "caused by and manifested in multiple problems [*and*] those problems have multiple origins" (Valentine, 2016, p.239) or in other words, the problems are contextually embedded. The example of youth disadvantage demonstrates this point. Youth disadvantage is often conceptualised as an accumulation of disadvantage, the various

effects of which are evidenced over the life course. This includes for example, how early environmental conditions affect adolescent cognitive/non-cognitive capabilities (Heckman, 2006, 2008, 2012); how access to education (Walsemann, Geronimus and Gee, 2008) and adverse childhood experiences (Bellis et al., 2013) can affect health and wellbeing in adulthood (Olsson, McGee, Nada-Raja and Williams, 2013) as well as overall life satisfaction (Layard et al., 2014). These studies highlight the 'multiplicity' of youth disadvantage through exploring the links between multiple causes and consequences throughout the life course. Being able to plan for, deliver and evaluate, not only the different components and processes of programmes aimed at tackling such contextually embedded social issues, but also capturing the changes in outcomes for beneficiaries, is a challenge in terms of programme evaluation. This is because it becomes difficult to demonstrate how a programme or service has contributed to changes in outcomes, outcomes which are caused by and manifested in many parts of a programme beneficiary's life.

Moreover, third sector organisations are often considered as well placed to tackle these challenging social issues because of their distinct knowledge of the communities and people to whom, and with whom, they deliver programmes and services (Macmillan, 2010; Lang and Hardwick, 2016), and thus their ability to design programmes and services to meet local needs (Hardwick, 2018), often using socially innovative solutions. As a result, many public services are now contracted to local third sector organisations (Hardwick, 2018). The importance of programme evaluation is therefore paramount in small TSOs in order to ensure that the programmes and services they provide are as successful as possible, and also so that we can learn about innovative ways of working with local communities.

It follows that 'social programming' has significance for small TSOs. Social programming relates to how social programmes are designed to contribute to social change (Shadish, Cook and Leviton, 1991). The design of social programmes to contribute to social change is achieved through understanding: how social programmes are structured and implemented; the external factors that influence programmes; and how programme change contributes to social change (ibid. 1991). These factors relate to programme implementation, the external context in which programmes are implemented, and how these two factors interact and are leveraged to contribute to social change. The distinct knowledge of change in small TSOs is therefore considered valuable in social programming efforts. The use of programme theory in programme evaluation is directly related to the concept of social programming whereby the development of programme theory makes explicit some of these assumptions regarding programme implementation and social change: this information is then used to evaluate the programme.

3. As a consequence of the dynamic internal and external organisational environments, and contextual embeddedness of the social issues targeted, the practice of programme evaluation within third sector organisations can often simultaneously take several forms and serve several purposes e.g. monitoring and accountability, programme improvement, and organisational learning (Carman, 2011). Organisational analysis theory[11] helps to explain in why programme evaluation is demanded, and in what ways, and how it serves such a wide range of purposes (Carman, 2011; Dahler-Larsen, 2012). For example, institutional organisation theory helps us to understand that evaluation is sometimes in social demand, e.g. through seeking

[11] Organisational analysis theory in this case refers to how different conceptualisations of the organisation can help to explain practice.

organisational legitimacy, even if there is little desire to use the evaluation within the organisation, meaning that evaluation processes are often disjointed and inconsistent (Dahler-Larsen, 2012). Power's notion of 'The Audit Society' supports these ideas that question the actual value, meaningfulness, and purpose of conducting programme evaluation activities (1997).

In practice, whilst evaluative activity in third sector organisations is shown to be dominated by accountability requirements and the search for legitimacy (Ellis and Gregory, 2009), there are increasing calls to support these organisations in ensuring their long-term sustainability through using the internal learning which programme evaluation activity can generate for the organisation (Ellis and Gregory, 2009; Inspiring Impact, 2013; Carman, 2011). Such learning can facilitate better cultures of evaluation where those practising evaluation, along with other stakeholders, are able to benefit from evaluation activity beyond any one specific evaluation (Preskill and Torres, 1999b). In short, it is acknowledged that these organisations are faced with challenging internal and external organisational environments, environments which constrain, demand and rely upon their ability to do evaluation.

The increasingly important role played by TSOs in local communities, and at a national level, means that the role of programme evaluation in such organisations has also gained interest, both in practice and within academic communities. The context of TSOs, particularly those which are small in size, has implications for the practice of programme evaluation in three ways. First, the dynamic internal and external organisational environments in which third sector organisations are situated present both a resource challenge for the capacity to conduct evaluation, as well as a challenging and often imbalanced accountability environment which places greater emphasis on economic-financial reporting rather than social value reporting. Second, some of the

social issues which TSOs are trying to tackle are often contextually embedded which makes social change, as a result of a programme or service, difficult to observe and report upon. Last, despite the challenging organisational environment and complexity of social issues targeted, there are calls for the practice of evaluation in third sector organisations to facilitate learning and cultures of evaluation that will ensure long-term organisational sustainability over and above short-term funding cycles (Ellis and Gregory, 2009; Inspiring Impact, 2013; Carman, 2011).

1.3.4 Programme theory in evaluation practice

The definition of research on evaluation provided earlier in this chapter states that a focus on evaluation methods and approaches was a particularly important focus for research on evaluation. Galport and Galport note that it is important for research on evaluation to pay attention to assessing tools, methods and approaches for conducting evaluation in order to understand the relative value of specific tools, methods, and approaches and under which conditions, or in which contexts, they are most likely to be appropriate and useful (2015). One of the challenges faced by those practising evaluation is the substantial choice of methodological approaches and tools available to them. Due to the lack of purposeful, systematic, and empirical consideration of these methods and approaches in the evaluation literature, it is often unclear what relative value methodologies or approaches can have over one another. However, in doing such research, evaluators can, as a result, “assess evaluation’s tools with the same rigor applied to other evaluands” (2015, p.27). This, in turn, allows better assessment of and comparison between evaluation methodologies and approaches based on how feasible they are to implement in different contexts, how valid the method or approach is at addressing evaluation needs, and how useful the method or approach is at generating the information or learning needed (Miller, 2010).

In this thesis I focus on a specific approach to evaluation called 'programme theory' (defined in the shaded box below). Broadly speaking, the use of programme theory in evaluation practice refers to the making explicit of the underlying assumptions about how and why programmes work to achieve some outcome(s) and using this information to support and guide programme evaluation in some way (Rogers and McDonald, 1999). The concept of programme theory therefore relates to the concept of social programming described in the previous section of this chapter, which highlights the importance of understanding how social programmes contribute to social change. Specifically, the use of programme theory involves the making explicit the links between programme/service activities (programme process or theory-of-action) and the outcomes which the programme is intended to influence (programme impact or theory-of-change).

DEFINITION

Programme theory is a specific approach used within the evaluation field to conduct evaluation. Programme theory is the explicit use of the assumptions about how and why programmes work to achieve some outcome(s), to structure, guide, and conduct programme evaluation in some way. Programme theory is often depicted in a diagram, commonly a box-an-arrow style diagram. Programme theory can therefore be considered as a method or approach falling under the umbrella of an evaluation theory.

The making explicit of the assumptions underlying how and why a programme works usually involves the development of a model, or visual representation, of the programme theory: this is normally in the form of a box-and-arrow type diagram. The linearity of box-and-arrow diagrams is, however, criticised and other methods have been developed that aim to better account for complexity and non-linearity in programme theory (Rosas, 2005; Dalkin et al., 2018). In any case, this model of programme theory will most likely include information

about programme process, i.e. the more operational aspects of the programme, and information on the programme impact, i.e. the short-, medium-, and long-term outcomes which the programme is attempting to change. Models of programme theory generally reflect the linkages between both programme process and programme impact. A simplified and generic example of a programme theory model is shown in Figure 1.

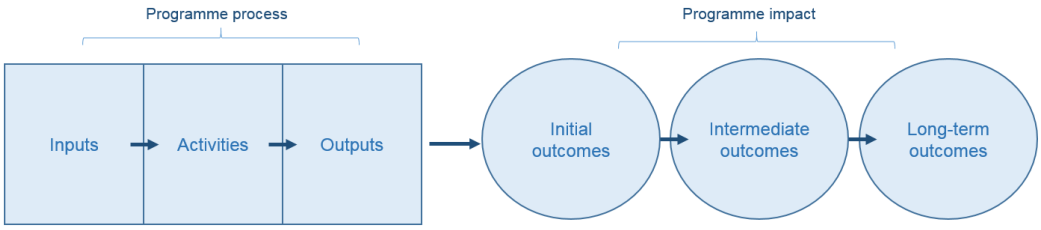


Figure 1 Generic example of programme theory model (adapted from Donaldson (2007))

The role of programme theory in programme evaluation has been defined in a number of ways. For example, a programme evaluation could 'test' a programme theory to understand whether it was the programme that actually brought about some change in outcomes (Chen and Rossi, 1983). Alternatively, an evaluation could use programme theory to consider which programmes or strategies would best contribute to social change in the first place (Weiss, 1995) or to generate knowledge about which mechanisms across different programmes contexts are most likely to contribute or improve some outcome (Pawson and Tilley, 1997). Programme theory can also be used to help identify and define appropriate outcomes and indicators to evaluate (McLaughlin and Jordan, 1999) and as part of the assessment of whether a programme is 'ready' to be evaluated, i.e. that it is properly defined and implemented (Wholey, 1987; Craig and Campbell, 2015).

There are several reasons why programme theory is an interesting focus for research on programme evaluation more generally and in the context of small

TSOs more specifically. First, since it was popularised in the 1980s, programme theory has increasingly appeared in evaluation journals, as evidenced by the graph in Figure 2. Moreover, it now forms the basis of several practical guidelines to conducting programme evaluation (Hatry, Van Houten, Plantz and Greenway, 1996; Evaluation Support Scotland, 1998; HM Treasury, 2018). There is now a plethora of different approaches to developing and using programme theory, signifying its popularity and interest, but also highlighting that the landscape of literature and guidance on programme can be confusing. These approaches include, but are not limited to, theory-driven evaluation, theory-based evaluation, theory-of-change evaluation, realist evaluation, logic analysis, contribution analysis, evaluability assessment, and logic modelling. I have summarised some of these approaches in the table in Appendix 1.

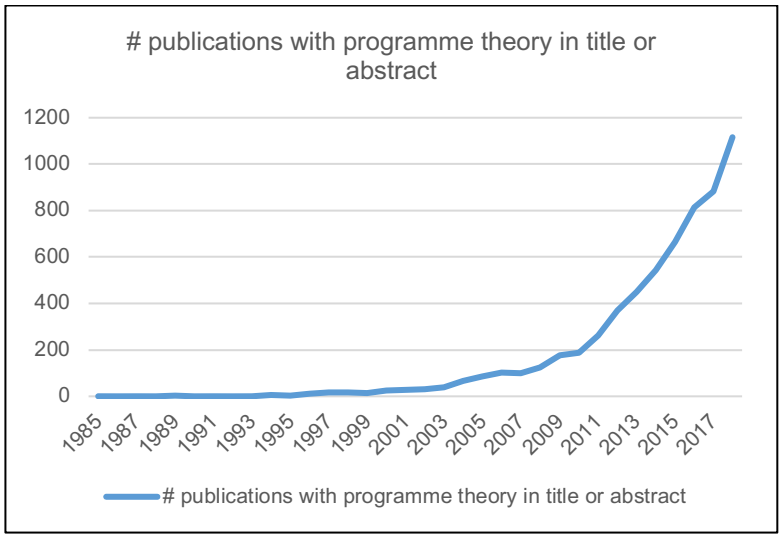


Figure 2 # mentions of programme theory in paper title or abstract - retrieved from Web of Science (August, 2019)

Moreover, in reviewing the literature on the use of programme theory in evaluation practice (discussed in more detail in Chapter 2), it is clear that the dominant perspective on evaluation practice reflected is not consistent with our understanding of programme evaluation practice in small TSOs for several reasons:

1. This literature is predominantly from the perspective of trained evaluators. However, those practising evaluation in small TSOs are more accurately defined as ‘accidental evaluators’ as they commonly have little or no formal evaluation training and conduct evaluation as just one part of their job. This has two implications for our understanding of the use of programme theory in programme evaluation practice:
    a. The literature on the use of programme theory in programme evaluation practice primarily focuses on methodological implementation of specific approaches to developing and using programme theory through tracking the application of methodological principles. However, in the case of ‘accidental evaluators’, it is more likely that their own implicit theories guide evaluation practice due to their lack of experience with evaluation theory. This means that whilst those practising evaluation in small TSOs might use programme theory in some form, it may not be appropriate, or feasible, to track the implementation of methodological principles in this setting.
    b. Moreover, in cases where there is little or no formal evaluation training, it is important to consider the capacity to do evaluation in terms of both evaluation skills and knowledge but also in terms of the ability to ‘think evaluatively’. Evaluative thinking is the thinking skills required to support the conduct of feasible and useful formal (systematic) evaluation activities. The literature reviewed in Chapter 2 focuses more explicitly on the use of programme theory in formal (systematic) evaluation activities rather than understanding *how* the development and use of programme theory facilitates evaluative thinking. Evaluative thinking is however important to consider in the case of small TSOs where evaluation capacity may be lower than in the case of trained evaluators.
2. The literature on the use of programme theory in evaluation practice takes little account of how contextual influences (e.g. the social, political

and organisational environments of those practising evaluation) influence the use of programme theory. As discussed in the preceding section on evaluation practice in small TSOs, the dynamic internal and external environments in which small TSOs are situated influences evaluation practice. As such, context is an important factor to consider in understanding the use of programme theory in small TSOs. Despite this, in the literature on the use of programme theory, it is the context of programmes themselves which has attracted more interest, rather than the context of programme evaluation practice.

1.4 Research aims and objectives

It follows that the overall research aim of this thesis is to:

Explore the role of the use of programme theory in programme evaluation practice in small third sector organisations

1.4.1 Research objectives

The objectives of addressing the overall research aim relate to both contributing to our understanding of the use of programme theory in the setting of small third sector organisations and the implications such understanding might have for evaluation practice.

1. Discuss the key differences between evaluation theory and evaluation practice with respect to the use of programme theory in small third sector organisations.
2. Empirically examine the role of the use of programme theory in programme evaluation practice in small third sector organisations in order to:
    a. Inform the theoretical understanding of the use of programme theory in small third sector organisations using a methodological approach which facilitates a dialogue between prescriptive and descriptive theory
    b. Consider the ways in which programme evaluation capacity and evaluation practices in small third sector organisations can

be improved, developed, and supported through the use of programme theory.

1.5 Research questions

The literature on the use of programme theory in evaluation practice is not consistent with our understanding of evaluation practice in small TSOs. As a result, the aim of this thesis is to explore the use of programme theory in the context of programme evaluation practice in such organisations. In order to explore the role of programme theory in small TSOs, this thesis addresses two research questions which focus on both descriptive and prescriptive aspects of the research problem:

RQ1: What is the current practice with respect to the use of programme theory in programme evaluation practice in small TSOs?

a. What is the context of programme evaluation practice in small TSOs with respect to the capacity to do evaluation, evaluation activities, and the use of evaluation?
b. In what forms, does the use of programme theory manifest itself within the context of programme evaluation practice in small TSOs?

RQ2: How can the development of programme theory improve future programme evaluation practices within small TSOs (particularly with respect to facilitating evaluative thinking)?

1.6 Summary of research approach

The research questions defined, based on the review of the literature on the use of programme theory in programme evaluation practice, are respectively descriptive and prescriptive in nature. In this way, the first research question aims to describe current practice and the second research question aims to prescribe how the use of programme theory can improve future programme evaluation practice. Taken together, by addressing these research questions simultaneously, this thesis can form the beginnings of informing a more effective dialogue between theory and practice in the context of programme

evaluation in small TSOs. It follows that I choose to adopt a multi-methodological framework, within which I use two different methods to address each research question individually. This multi-methodological framework is underpinned by a pragmatic philosophy that shifts methodological choices away from metaphysical questions about the nature of reality to a focus on how the beliefs of researchers inform their actions through the process of inquiry (research) (Morgan, 2014). Whilst this view does not reject the existence of some reality, it maintains that our ability to access that reality in research is limited and that rather than contributing to some explicit 'knowledge', what we claim to generate in research is related more to 'warranted assertions'.

To address RQ1, I use qualitative interviews (semi-structured interviews) to elicit the perspectives of those practising evaluation in small TSOs. By doing so I elicit implicit theories of evaluation practice with particular attention to the use of programme theory in the context of programme evaluation practice. In this study, I conducted 23 semi-structured interviews, primarily with those practising evaluation in small TSOs, but also with participants from organisations who fund small TSOs, and organisations who support evaluation practice in small TSOs. I analyse the interview data using thematic analysis. Findings presented relate to programme evaluation practice and the contextual influences on this practice, as well as the current use of programme theory.

To address RQ2, I use action research to understand how developing programme theory can improve evaluation practice in small TSOs. This involved implementing an action research intervention (the development of programme theory) in two organisations. This work consisted of working closely with both organisations, over a period of roughly 6 months, to identify their evaluation needs, to define an approach to developing programmes that met those needs, and conducting facilitated discussions and workshops to

understand how the development of programme theory met those evaluation needs. Learning presented from this study relates to the role of programme theory in the two organisations within which programme theory was developed.

Using a multimethodological framework requires a process of 'theoretical integration' to triangulate the findings. To do so, I use the discussion chapter of the thesis to bring together the learning generated in each study to consider and reflect more critically on the use of programme theory in the context of programme evaluation practice in small TSOs.

1.7 Thesis contributions

The learning generated from this thesis makes an academic contribution through improving our understanding of, and better conceptualising, the use of programme theory in small TSOs, both in terms of the use of programme theory in current programme evaluation practice and how its use can improve future programme evaluation practice.

In particular, the use of a practitioner-oriented perspective to consider the use of programme theory in this context, contrasts with the dominant perspective in the literature on the use of programme theory, which focuses primarily on methodological implementation. This contribution also highlights the value of adopting a practitioner-oriented perspective in research on evaluation more generally. The additional understanding that can be generated from a practitioner-oriented perspective supports learning about how evaluation methods and approaches are used in programme evaluation practice, and how they can improve that practice.

Methodologically, the multi-methodological framework used enables this thesis to generate descriptive understanding about the current use of

programme theory in small TSOs, but also to consider more prescriptively how the use of programme theory can improve programme evaluation practice in the future. In addressing both descriptive and prescriptive components of the research problem simultaneously, this thesis develops an effective dialogue between descriptive and prescriptive aspects of the research problem. In particular, the use of action research should be highlighted as a particular contribution and its use in the field of research on evaluation advocated.

Together, these contributions have several implications for evaluation practice in small TSOs. First, this thesis provides a clearer sense of the needs of those practising evaluation in small TSOs, particularly in terms of providing better resources for developing and using programme theory. Second, this thesis develops our understanding of how to support organisations in the development of programme theory in this context. Last, this thesis has implications for the teaching of evaluation in terms of ensuring that the rhetoric around evaluation is not solely dominated by a methodological and trained evaluator perspective. Rather, we need a rhetoric that acknowledges the diverse landscape of programme evaluation practice, a landscape in which not all evaluators are formally trained in evaluation, and who often work in contexts which present additional challenges for, and influences on, conducting programme evaluation.

1.8 Thesis structure

Chapter 2 – Literature review

In Chapter 2, I first review the literature on programme evaluation practice. I argue that the various ways in which evaluation practice is defined shapes the focus and methodological approach of research on evaluation practice. The second half of the literature review focuses on the literature on the use of programme theory in evaluation practice. This chapter serves to highlight and define the research problem and specific research questions addressed in this thesis.

Chapter 3 – Methodological framework

Given the research questions identified in Chapter 2, Chapter 3 defines and justifies the multi-methodological framework adopted to address the research questions. I first justify the multi-methodological framework before defining in more detail the two specific approaches (qualitative interviews and action research) adopted within this methodological framework to answer each research question respectively.

Chapter 4 – Study 1 findings

Chapter 4 presents the findings of Study 1, which is the qualitative study of current programme evaluation practice in small TSOs. This involves presenting the thematic analysis of semi-structured interview data. This chapter first presents findings relating to programme evaluation practice in context in small TSOs. Second, this chapter presents findings relating to the use of programme theory in this sample.

Chapter 5 – Study 2 findings

Chapter 5 presents the findings of Study 2, which is the action research study focusing on how the development of programme theory can improve future programme evaluation practice in small TSOs. This chapter respectively presents the two organisations with whom I worked, the programmes which they deliver, the specific evaluation challenges faced within each organisation, the approach to implementing the action research intervention, and the learning derived in each instance of developing programme theory. Lastly, this chapter provides some key learning points as they findings emerging from both organisations.

Chapter 6 – Discussion

Chapter 6 brings together the findings from Study 1 and Study 2 to consider the research questions defined earlier in the thesis. I consider the findings of

Study 1 in light of previous research on programme evaluation practice in small TSOs. The findings from Study 2 generate emergent learning about how the development of programme theory can improve evaluation practice in small TSOs. As such, I introduce some ideas which help to explain how the development of programme theory can improve programme evaluation practice. I consider this discussion as the beginnings of a conceptual framework to understand the use and value of developing programme theory in this setting. This research focuses on the use of programme theory in small TSOs, but the conceptual framework developed also has implications for the role of programme theory in other types of organisations or other settings in which programme evaluation is conducted.

Chapter 7 – Concluding the thesis

In Chapter 7, I consider the theoretical, methodological, and practical implications of this research. I also consider the limitations of this thesis and its potential for publication. Finally, I include my own reflections on conducting this research.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

In the introductory chapter to this thesis, I identify the need for research on programme evaluation practice to address the gaps between prescriptive and descriptive evaluation theory, and to inform better dialogue between theory and practice in the evaluation field. Schwandt argues that without learning about practice, evaluation theory is abstract and lacks any concrete meaning for those practising evaluation (2014). I define research on evaluation practice as any purposeful, systematic, and empirical study on evaluation processes, products, or evaluation theories, methods, or practices (Coryn et al., 2016). Defining research on evaluation practice in this way stands in contrast to anecdotal reports of practice which are the most commonly found types of studies in the evaluation literature (Christie, 2003c; Vallin et al., 2015). These kinds of studies are largely post-hoc and descriptive reflections and reports of evaluation activities and findings that lack an obvious framework to help the reader to summarise findings across studies with respect to evaluation theory (Vallin et al., 2015; Coryn et al., 2016). In such cases, "*findings are more emic than etic and more idiographic than nomothetic"* (Vallin et al., 2015, p.14). In other words, findings produced by anecdotal or single-case reports of practice are not necessarily relevant beyond the specific case of evaluation.

The goal of research on evaluation practice is to develop theory that supports feasible practices that support those practising evaluation and ultimately to inform and support the implementation and evaluation of programmes to target the social problems facing society (Shadish, Cook and Leviton, 1991; Mark, Henry and Julnes, 2000) by informing a dialogue between prescriptive and descriptive theory. However, an important question to address with respect to research on evaluation practice is what is meant by evaluation practice? In other words, when we refer to evaluation practice, to what are we referring? What seems to be an intuitive answer, that defining evaluation practice as the 'doing of evaluation', is, I argue too broad a definition to conduct systematic and purposeful research on evaluation practice.

The focus of this thesis is on the use of programme theory in the context of programme evaluation practice in small TSOs. However, I first use this chapter to review the literature conceptualising, or defining, evaluation practice. I argue that how we conceptualise evaluation practice will affect that which research on evaluation focuses on as well as how research on evaluation is conducted. I use the latter half of this chapter to lay out how the literature on programme theory has conceptualised the use of programme theory in evaluation practice.

Therefore, in this chapter, I summarise the literature that helps us to understand the various ways in which evaluation practice is defined and conceptualised in the literature. I then use this understanding as a framework through which to consider the conceptualisation of the use of programme theory in practice in the latter half of the chapter. In doing so, I define the research problem that the empirical component of the thesis will address. This section of the review aims to highlight that different conceptualisations of evaluation practice have distinct value depending on the perspective which the researcher adopts, the research question addressed, and the context of evaluation practice context being studied.

2.2 Literature review approach

The first section of this literature review focusses on the literature defining evaluation practice. Given the vastness of this literature, I read widely to identify some key concepts that would help me understand practice as a concept. This led me to areas of the literature that I could explore in more detail. This included, evaluation capacity (including different types of evaluators), evaluation activities (including evaluative thinking), the use of evaluation, and evaluation context. These guiding concepts provided me with a more structured framework through which to search the literature, initially

beginning with seminal papers in each area and using a snowball approach to identify key papers which had cited those seminal papers.

In terms of the literature on programme theory, I recognised at an early stage that there was a huge literature reporting on specific applications of programme theory. I therefore utilised some key search terms to structure my search within this wider body of literature. These search terms related to the different approaches to using programme theory, e.g. program* theory, theory-based evaluation, theory-of-change, realist evaluation, contribution analysis, logic analysis, logic model, evaluability assessment.

I did not wish to define this literature review as a systematic literature review as the purpose was to identify the gaps in the research, rather than address a specific research question (Greenhalgh, Thorne and Malterud, 2018). I sought to utilise the literature review to be critical and emphasise a different lens through which to look at the use of programme theory in evaluation practice.

2.3 Defining evaluation practice

Evaluation practice is often cited as the 'doing of' evaluation, or what evaluators do when they conduct evaluation (Rog, 2015). However, for a variety of reasons, this conceptualisation of evaluation practice is not helpful in terms of defining empirical research on programme evaluation practice that is purposeful and systematic in nature. First, it does not address *who* is practising evaluation i.e. who is the evaluation practitioner. Second, it does not indicate *what* exactly is being practised, i.e. the evaluation activities conducted. Third, it does not conceptualise *how* evaluation is being practised. Last, it does not address the contextual influences on *why* evaluation is being conducted in the way that it is. In the following sections, I discuss each of these questions in turn in order to lay out the different ways in which programme evaluation practice can be conceptualised. Specifically, I discuss the

implications of how we conceptualise evaluation practice on conducting research on evaluation. Each definition can have distinct influences on the focus of, and/or the methodological approach taken, to conduct research on evaluation practice.

I argue that considering these questions regarding the who, what, how, and why of evaluation practice provides a more refined conceptualisation of evaluation practice that can be used to conduct research on evaluation. Not only can this help to identify gaps in our knowledge about a given evaluation theory but it can also influence how we conduct research on programme evaluation practice. In terms of this research, a more refined conceptualisation of programme evaluation practice means I can address the research problem of the use of programme theory in programme evaluation practice in small TSOs in a more nuanced and valuable way – in that way, research can contribute to, and inform a dialogue between, both theory and practice.

2.3.1 Who is practising evaluation?

A natural starting point in conceptualising evaluation practice is the question of who is actually conducting evaluation, i.e. who is the evaluation practitioner? The literature on evaluation practice points to two different types of people who practice evaluation:

1. One is the trained evaluator or applied social scientist. This evaluation practitioner has formal training in evaluation theory and methods and/or is a trained social scientist i.e. trained in social research methodology. Their skills lie in their ability to apply evaluation theory, methods, and approaches to the evaluation of programmes, projects, and interventions.
2. A second type of person who practises evaluation is characterised as an ‘accidental evaluator’ (King, 2003) i.e. someone who is not formally trained in evaluation and does not conduct evaluation as a primary

component of their job (Alkin, 2003; Rog, 2015). Few of those practising evaluation in this case have received formal evaluation training because they are often in roles in some other field (Christie, 2003a) e.g. social work, youth work, or health.

2.3.1.1 Implications for research on programme evaluation

	Accidental evaluator	**Trained evaluator**
Capacity to do evaluation (evaluation skills and knowledge)	Low	High
Substantive knowledge vs generalist knowledge	High substantive Low technical	Low/high substantive High technical
Extent to which programme evaluation practices draw on evaluation theory	-ve	+ve
What guides programme evaluation practice?	Implicit theories of programme evaluation	Programme evaluation theory/social science methodology

Table 1 Defining who is practising evaluation: summary of implications for research on evaluation

How we conceptualise who is practising evaluation has implications for our understanding of the use of evaluation theory in practice and the conduct of research on evaluation practice. These implications are summarised in Table 1. First, Table 1 highlights the issue of evaluation capacity i.e. the ability (in terms of individual skills, behaviour, knowledge, and attitudes as well as organisational support structures) to do evaluation (Labin et al., 2012; Cousins, Goh, Elliott and Bourgeois, 2014), as well as the level of substantive, or topic, knowledge. Second and relatedly, it raises the question of the extent to which evaluation practices draw on evaluation theory. Naturally, this will depend on the extent of experience with evaluation theories, methods, and approaches through training, support, and involvement in evaluation (Labin et al., 2012). Last, if there is little to no training in evaluation theories, methods, and

approaches, the question remains about what guides evaluation practices in such situations?

First, let us consider evaluation capacity, or the capacity to do evaluation. Importantly, building the capacity of practitioners to do evaluation is reported to have links to supporting the utility of both the process of doing evaluation as well as the findings or outcomes of evaluation activities (Johnson et al., 2009; Cousins, Goh, Clark and Lee, 2004; Cousins et al., 2014; Labin et al., 2012). As a result, much attention in the evaluation literature has been paid to 'evaluation capacity building' (Preskill and Boyle, 2008), as an intentional process to increase the motivation, knowledge and evaluation skills of individuals and therein facilitate a group or organisation's ability to conduct and use evaluation (Labin et al., 2012, p.308) particularly within settings where those involved in the evaluation have little or no formal evaluation training.

Similarly, depending on whether the evaluators are accidental evaluators or trained evaluators, they may have differing levels of substantive knowledge (or subject knowledge) of the programme, and generalist technical knowledge (or evaluation knowledge and skills) (Rog, 2015). This raises the question of the extent to which formally trained evaluators have sufficient substantive expertise of that which they are evaluating and, on the other hand, the extent to which 'accidental evaluators' have sufficient evaluation expertise to conduct programme evaluation activities. In terms of research on programme evaluation practice, the value of having subject area expertise or evaluation expertise is important to consider. For example, one might assume that an accidental evaluator has high substantive knowledge of the programmes and services that they deliver, with low level knowledge about how to conduct programme evaluation. A trained evaluator, who works internally within an organisation, might have both substantive expertise paired with formal programme evaluation training. A trained evaluator, who is external to the organisation, may have expertise in evaluation theory, methods, or

approaches, but less extensive substantive knowledge i.e. they may have more generalist evaluation methodology skills. In this sense, the accidental-trained evaluator conceptualisation does not necessarily reflect a dichotomy, rather a spectrum of differing levels of substantive and technical expertise amongst those practising evaluation. Nevertheless, perhaps it is too often assumed that evaluation expertise (i.e. evaluation knowledge and skills) is the central determinant of capacity to do evaluation when perhaps it is also important to consider the value of substantive knowledge.

Second, if we define those practising evaluation as trained evaluators or social scientists, we might expect them to have good knowledge of and the skills relevant to evaluation theories, methods, and approaches, as well as the ability to apply these in a way which is considered appropriate for a given evaluation. Research on the evaluation practice from the perspective of the trained evaluator will assume that evaluation is a core component of evaluation practitioners' jobs. Evaluators could be both external and/or internal to the organisations within which they are conducting evaluation. Many studies of evaluation practice come from the perspective of trained evaluators, potentially as a consequence of being motivated and able to formally publish empirical examples of the evaluations they have conducted. An example of a study focussing on trained evaluators is Chouinard and colleagues (2017) who focus on evaluation practice from the perspective of student evaluators moving from the 'classroom' to the practice of evaluation in the field.

Likewise, if we define those practising evaluation as 'accidental evaluators' we might expect, given a lack of formal training, a more limited capacity in terms of knowledge, skills, and application of evaluation theories, methods, and approaches. In practice, Alkin notes that it is often difficult, in cases of accidental evaluators, to set training goals due to the 'in transit' nature of the workforce in such settings (2003). An empirical study evidences the gap between evaluation theory and practice (Christie, 2003c): analyses point to

greater coherence between external evaluators' practice and evaluation theory than is the case with 'accidental evaluators' practices (Christie, 2003c). However, it is also worth noting that the same study also found gaps between evaluation theory and external evaluators' practices as well. If those who are practising programme evaluation are 'accidental evaluators', research on programme evaluation in such settings should be mindful of the lack of formal training in programme evaluation. By acknowledging a lack of formal training, research can better account for a wider set of influences on, and experiences of, programme evaluation practice.

A final, and related, implication of viewing evaluation practice from the perspective of who is doing evaluation is the question of when there is little, or no, formal evaluation training, what then guides evaluation practice? Some argue that evaluation conducted by 'accidental evaluators' is based on informal, implicit, or 'folk' theories derived from experiences and ideas about evaluation, and the values of practitioners, rather than formal evaluation theory, methods, or approaches (Christie, 2003a). Christie highlights that such evaluators, acting on conceptions and assumptions about evaluation, are implementing their implicit theories about evaluation (2003b) i.e. "*how everyday practitioners form notions about evaluation and then use them to guide their work*" (2003b, p.92). To further emphasise this point, Christie's (2003c) empirical quantitative study on the match between theorists' ideas and practitioners' practice argues that internal evaluators are in many cases programme administrators who conduct evaluation as only one part of their job, and so argues that theory is not a prerequisite to evaluation practice in all cases (2003c). What Christie implies is that implicit theories may not therefore mirror prescriptive programme evaluation theory at all. Studies of how evaluation practitioners perceive their own work are therefore important to developing more realistic and appropriate evaluation theory, methods, and approaches (Smith, 1993), particularly in settings where the evaluation practitioner is not formally trained in evaluation.

2.3.2 What evaluation activities are being practised?

A second factor to consider when conceptualising programme evaluation practice is the question of what is actually being practised i.e. what are evaluation practitioners actually doing when they do programme evaluation? In what follows, I discuss 'formal evaluation activities' and 'evaluative thinking' as two distinct but related aspects of what is being practised in programme evaluation.

2.3.2.1 Formal (systematic) programme evaluation activities

DEFINITION

Formal programme evaluation activities are *systematic* inquiries that can use a variety of methods and approaches, and can be conducted at any stage in the programme lifecycle, to describe or explain the programme in terms of its implementation, effects, and/or social implications, to help evaluators and evaluation stakeholders make sense of a programme (Mark, Henry and Julnes, 2000).

The most obvious starting point, with respect to what is being practised, is what I refer to as formal (systematic) evaluation activities. Formal evaluation activities are evaluation activities that involve conducting *systematic* inquiry at any stage in the programme life cycle to describe or explain the programme in terms of its implementation, effects, and/or social implications; such systematic inquiry can be conducted using a variety of methods and approaches (Mark, Henry and Julnes, 2000, p.3). Mark, Henry and Julnes state that these formal evaluation activities enable stakeholders to better "select, oversee, improve, and make sense of social programmes and policies" (Mark, Henry and Julnes, 2000, p.3).

There are several factors to consider conducting formal evaluation activities: for what purposes is the evaluator conducting formal evaluation activities; when will the evaluation activities be carried out during the lifecycle of the programme, i.e. will it be a formative or a summative evaluation; and what methods and approaches will be used to carry out the evaluation activities (Mark, Henry and Julnes, 1999, 2000; Langbein, 2012). I discuss these questions in the following paragraphs.

Evaluation can serve several distinct purposes which taken individually will impact how one goes about conducting systematic inquiry, or formal evaluation activities. Within their framework on conducting evaluation, Mark, Henry, and Julnes outline four purposes which formal evaluation activities might serve (1999, 2000):

1. Assessment of merit or worth of the programme, i.e. is the programme working well?
2. For programme and organisational improvement, i.e. what can we do to improve the programme?
3. For oversight and compliance, i.e. are we doing what we said we would do in the way we said we would do it?
4. For knowledge development, i.e. what can we learn about this programme or interventions that has implications for the wider population?

The purpose, or purposes, which formal evaluation activities serve then dictates the approach to systematic inquiry taken, i.e. the focus of the evaluation, the evaluation's methodological approach, and subsequently, how the formal (systematic) evaluation is carried out. The substantive focus of the evaluation is that which the evaluation activities will seek to understand, sometimes known as the evaluand. Commonly, the substantive focus is on the implementation of the programme and/or the outputs and outcomes of the programme for the target population (Langbein, 2012). However, other

examples of different factors which can be studied in formal evaluation activities are the needs of programme beneficiaries, programme inputs such as budgets and staffing, contextual variables influencing the programme, and the costs and benefits of the programme (Shadish, Cook and Leviton, 1991).

The methodological approach to conducting a formal evaluation reflects what one wishes to understand about the implementation or outcomes. Does the evaluation seek to describe or explain the implementation or changes in outcomes? A descriptive evaluation might seek to describe the target population or describe the implementation of a programme whereas a causal, or explanatory, evaluation aims to assess whether the programme brought about desired outcomes (Langbein, 2012). Other evaluations may focus on performance monitoring and measurement (Scheirer and Newcomer, 2001). In any case, the substantive and methodological focus of the evaluation will have implications for how an evaluator carries out the evaluation in terms of the evaluation questions addressed, the methods used to collect and analyse data, the role of the evaluator, and how the findings of the programme evaluation activities will be used (Shadish, Cook and Leviton, 1991).

Inevitably however, the ability to conduct formal evaluation activities successfully, as described in the preceding sections, will ultimately depend on the capacity to do evaluation of those practising evaluation, in terms of skills and knowledge of evaluation theory, but also in terms of the ability to apply such skills and knowledge through ‘evaluative thinking’.

2.3.2.2 Evaluative thinking

DEFINITION

Evaluative thinking refers to the thinking skills (critical, creative, inferential, and practical thinking skills) required to frame, design, and carry out formal evaluation activities that are feasible and useful to evaluation stakeholders.

Formal (systematic) evaluation activities are often at the forefront of an evaluator's mind. However, an important skill and process underlying the conduct of formal evaluation activities is 'evaluative thinking'. In addition to the skills and knowledge of evaluation theory, evaluative thinking supports the conduct of formal evaluation activities and the subsequent use of evaluation findings (Preskill and Russ-Eft, 2016). In the previous section on 'who is practising evaluation', I referred to the capacity to do evaluation, but I referred only to the knowledge and skills to conduct evaluation, knowledge and skills acquired through formal evaluation training. However, an important part of this capacity to do evaluation is 'evaluative thinking'. Evaluative thinking is not necessarily a precursor to formal evaluation activities, in the way that evaluation theory, skills and knowledge are, rather it is something that underpins the whole task of conducting formal evaluation activities and subsequently making use of the findings of the evaluation.

Being able to 'think evaluatively' is associated with the ability of evaluation practitioners to conduct appropriate, feasible and useful programme evaluation activities, particularly in terms of developing the critical thinking skills required to conduct evaluation (Preskill and Torres, 1999a; Preskill, 2014; Patton, 2011; Buckley, Archibald, Hargraves and Trochim, 2015; Patton, 2018). However, some argue that the association between evaluative thinking and critical thinking is too narrow and does not accurately reflect all that evaluative thinking entails. To conceptualise evaluative thinking more broadly, a series of authors have looked at both the historical development of the ideas

supporting evaluative thinking and current operationalisations of the concept in empirical research (Vo, Schreiber and Martin, 2018; Patton, 2018) to arrive at a more comprehensive and accurate definition of evaluative thinking.

Historically, evaluative thinking can be linked to several concepts outside the field of evaluation that see critical thinking as a process and an outcome for both learning and democracy (Patton, 2018). In this case, evaluation is seen as a democratic *process* which is inclusive of the voices of stakeholders and which supports dialogue and deliberation (House, cited in: Patton, 2018) through empowering and encouraging stakeholders to engage in critical thinking during the process of evaluation (Patton, 2018, 2002; Fetterman, 1994). As an *outcome,* evaluative thinking is therefore associated with enabling both evaluator and stakeholder to understand, apply, analyse, synthesise, and evaluate information at hand (Bloom, cited in: Patton, 2018). Whether defined as a process or outcome, the historical ideas behind the concept of evaluative thinking see its role in terms of generating useful learning, while also facilitating fair evaluation processes.

The concept of evaluative thinking as thinking critically has also been developed within the field of evaluation. Weiss, for example, draws from the concept of reflective practice (Schön, 1991) in helping evaluation practitioners reflect on their evaluation practice to think more critically about why the programme they are evaluating is operating as it does (Weiss, 1998). Extending the ideas of critical reflection is the concept of *process use* in evaluation. Process use (and evaluative thinking) are concepts which highlight the difference between 'learning how to learn' and learning something about an evaluand (Patton, 2008, 1998), where "learning how to think evaluatively is learning how to learn and think critically … [and] those who become involved in an evaluation learn by doing" (Patton, 2008). Process use, according to Patton, is an outcome of being involved in evaluation. Therefore the association between evaluative thinking and process use reflects evaluative

thinking as a more enduring skill, extending beyond simply making use of the findings of any individual evaluation (Patton, 2018).

Whilst the associations with critical thinking are clear, evaluative thinking can be conceptualised more broadly. Patton argues that evaluative thinking involves both the process and outcome being able to be creative and democratic in one's approach to designing and conducting formal evaluation activities, to reason well with evidence and build a persuasive argument for what the evidence shows, and to consider the real-world implications of evaluation activities (Patton, 2018). In Patton's words, evaluative thinking is the ability of an evaluator to employ critical thinking alongside 'creative thinking, inferential thinking, and practical thinking' to enable them to better and more rigorously define, frame, design, and carry out formal evaluation activities that are both contextually and ethically sensitive, and which generate findings that are useful in light of their strengths and weaknesses (Patton, 2018). An example of a practical use of evaluative thinking, using this definition, is in making boundary judgements in evaluation i.e. judgements "about what aspects of a situation are and ought to be part of the picture we create of what is being studied and evaluated" (Schwandt, 2018, p.131). Evaluative thinking is therefore more than simply critical thinking, it is about how to create a picture of that which is being evaluated through critical, creative, and reasoned inference alongside practical judgements.

Whilst the preceding discussion informs us of what evaluative thinking involves, it does not explain in detail how it occurs. Anne Vo and colleagues conduct a systematic review of literature related to evaluative thinking to better conceptualise how evaluative thinking occurs. The resulting findings of 220 records highlight four core domains of evaluative thinking in a conceptual model (shown in Figure 3): these are values and valuing, cognition, and application (Vo, Schreiber and Martin, 2018). This conceptual model implies that there is a mutually informing relationship between both the evaluator's socio-cultural values and evaluation-specific values, and the process of making value judgements about an evaluand which are intertwined by the cognition and application of evaluative thinking in the context of evaluation.

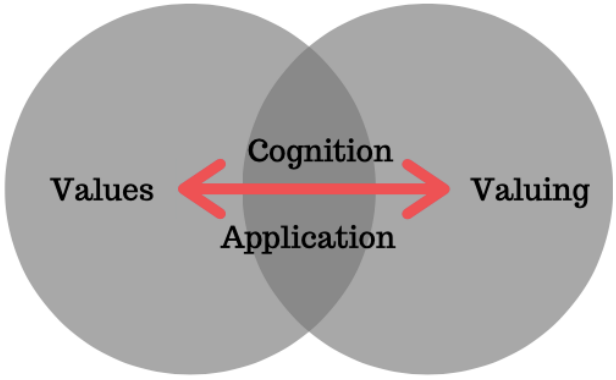


Figure 3 Model of evaluative thinking - adapted from Vo, Schreiber, and Martin (2018)

Central to this model are cognition and application which are the processes through which evaluative thinking occurs. Cognition refers to the technical aspects of evaluative thinking, at the individual and organisation level (Schwandt, 2018), that enable individuals and organisations to use critical thinking skills to make judgements based on evidence-based, systematic, and reasoned arguments in evaluation (Vo, Schreiber and Martin, 2018). These technical aspects of evaluative thinking are influenced by the application domain where 'application' reflects the actual doing of evaluation *in context* where the context specific challenges might affect the cognitive aspects of evaluative thinking. These challenges include evaluation capacity and the ability to use evaluation findings (Vo, Schreiber and Martin, 2018) where

evaluators need to have the skills and abilities to design, conduct, and use evaluation findings in ways which are useful *in context*. Overall, this model of evaluative thinking summarises much of the preceding discussion on what evaluative thinking is. Figure 3 illustrates that evaluative thinking is a process through which socio-contextual, personal, and evaluation-specific values are mutually linked to the process of making reasoned value judgements about an evaluand (e.g. a programme) through both cognition and application.

Linking evaluative thinking and formal evaluation activities

It is important to note that formal (systematic) programme evaluation activities and evaluative thinking are not mutually exclusive activities. This is reflected in the model developed by Vo and colleagues in Figure 3 (2018). Rather, evaluative thinking refers to the thinking skills required to conduct feasible and useful formal (systematic) evaluation activities. It follows that stronger evaluative thinking skills are aligned with the conduct of formal (systematic) evaluation activities that better meet the needs of evaluation practitioners and stakeholders.

2.3.2.3 Implications for research on programme evaluation

Research on programme evaluation, focussing on the use of a given evaluation theory in formal (systematic) programme evaluation activities, seeks to understand how that evaluation theory has been used to conduct systematic inquiry to describe or explain some aspect of the programme. For example, research on evaluation might seek to understand how a specific method or approach had been used to describe programme implementation. However, as discussed above, evaluative thinking refers to the thinking skills required to define, frame, design, and carry out formal programme evaluation activities in the first place, which provides a slightly different empirical focus for research on evaluation theory (methods or approaches). This focus would entail thinking more about how evaluation theory helps an evaluator to do evaluation, rather than what it helps them to do.

It follows that in many programme evaluation contexts, a consideration of evaluative thinking, in research on evaluation, may be a useful focus given the links between evaluative thinking and the conduct of formal (systematic) programme evaluation activities. An example of where a focus on evaluative thinking might be particularly useful is in programme evaluation contexts where those conducting evaluation have no formal training in evaluation and therefore might struggle to define, frame, design, and carry out formal programme evaluation activities. Understanding how, in such cases, a given evaluation theory either supports, or fails to support, evaluative thinking, can improve our understanding of how to better support the use of that evaluation theory. Research on programme evaluation, from the perspective of evaluative thinking, would seek to understand how the use of an evaluation theory enables those practising evaluation to critically reflect on their evaluation practice and, in turn, support the execution, design, conduct of, and use in formal (systematic) evaluation activities in contextually and ethically sensitive ways, whilst generating findings which are useful for evaluation stakeholders.

2.3.3 How is evaluation being practised?

In conceptualising the role of evaluation theory in practice, a third factor to consider is the question of *how* evaluation is being practised, i.e. how programme evaluation happens. In this sense, there are two ways to conceptualise how evaluation is being practised. The first conceptualisation of how evaluation is being practised is defined as instrumental, or technical, rationality. Instrumental rationality views evaluation practice as the site for the application of theory, method, or approach. The second conceptualisation of how evaluation is being practised views practice from the practitioner perspective: the use of language, behaviour, judgement, perceptions and experiences that forms implicit theories of evaluation, which then influence evaluation practice.

2.3.3.1 Instrumental rationality

The first conceptualisation of how evaluation is practised is referred to as instrumental rationality (Dahler-Larsen, 2012). This definition focuses on two things: the instrumental nature of evaluation practice (focus on evaluation as a tool, method, or technology), and the idea of rationality, that evaluation practice is, or should be, theory enacted (rather than an abstract process of inquiry) (Schwandt, 1989; Dahler-Larsen, 2014). Instrumental rationality therefore views evaluation practice as the site for the application of social science and in turn, accepts the ideas underpinning the notion of evidence-based practice (Schwandt, 2005). This view emphasises evaluator expertise, that the evaluator's application of social science will provide scientifically valid information about 'what works' (Schwandt, 2005). Through this perspective, evaluation serves an objective function with a focus on instrumental use for decision-making and control (Schwandt, 1989). In the wider management literature, practitioner-scholars define these ideas whereby "professional activity consists in instrumental problem solving made rigorous by the application of scientific theory and technique" (Schön, 1991, p.21). This view of how evaluation practice ultimately sees the relationship between theory and practice as somewhat top-down.

2.3.3.2 A practitioner-oriented perspective

The notion of instrumental rationality contrasts with a practitioner-oriented perspective and focusses less on the application of theory or method and more on the language, behaviour, and judgement of those practising evaluation, and thus views evaluation as a process of inquiry that occurs within and between individuals (Schwandt, 1989), a process that involves individual and collective efforts to make sense of the evaluation process and its findings (Mark, Henry and Julnes, 2000). Scholars, and in particular Thomas Schwandt, oppose the idea that practice is, or simply should be, theory enacted: he argues that practice cannot be 'evidence-based' as such, rather that practice 'on the ground' should inform theory. My earlier discussion of implicit theories of practice emphasises a similar idea, that evaluation practice is often not guided

by evaluation theory, methods, or approaches at all, placing more emphasis on the notion of practitioner judgement and decision-making. This practice focussed view therefore sees evaluation as a social practice that is morally and contextually influenced rather than an objective scientific undertaking (Schwandt, 2005) that can be studied using traditional positivist social science methodologies, focused on the generation of objective knowledge about practice .

In this way, Schwandt conceptualises practice, in his words as 'the practical', as a way of being and knowing that is always in the state of being accomplished (2003, p.354). He conceptualises 'the practical' according to several features of practice that move away from viewing evaluation practice as the application of theory. He argues that research on programme evaluation practice, should thus focus on 'the practical' through acknowledging that:

1. We make decisions and act in our lives whilst dealing with the uncertainties of everyday life. We make decisions and take actions not based on rational questions of what to do, but based on value-based questions such as what should I do, or what is desirable? There are not definitive or objective answers to these questions within this view on 'the practical';
2. We do not relate ourselves to each other and our surroundings through using scientific theory, rather we relate to these things in the way that we talk and act in everyday life;
3. 'The practical' refers to the fact that the world that practitioners operate in is real, linguistic, and contextual in that in order to arrive at decisions, they must exercise judgement rather than pure rationality and objectivity;
4. The process of reasoning in 'the practical' is not characterised by objectivity, rather it is a kind of practical reasoning that is sensitive to context, and involves interests, values, and emotions;

5. ‘The practical’ does not mean that evaluation should aim to claim ‘full intellectual control’ over a programme, rather, given that programmes are dynamic, that the goal of programme evaluation should be to continually re-define what one thinks as one continues to operate in ‘the practical’.

2.3.3.3 Implications for research on programme evaluation

Empirical inquiry adopting a view of instrumental rationality to understand the use of evaluation theory, methods, or approaches would seek to understand how the specific principles or standards of a theory, method, or approach had been applied in practice. As a dominant epistemology of practice, instrumental rationality means that our approach to conducting research on evaluation practice has been focused on the application of social science methods to study practice and, as a result, has a top-down focus on evidence (or theory) and its application (Schön, 1991). It therefore normatively perceives practice to be ‘evidence-based’, specialised, firmly bound, scientific, and standardised (Schön, 1991). With that, the narrative around how evaluation is practised and subsequently studied is focussed on methodology and specifically, how to devise and implement a set of procedures (Schwandt, 1989). Research on evaluation adopting this perspective would therefore have a top-down perspective on the conduct of research on programme evaluation where the researcher would seek to identify the principles or standards of a theory, method, approach and use those to track how they were being applied in practice.

Empirical inquiry into the use of evaluation theory, methods, and approaches that adopts Schwandt’s view on practice, i.e. what I term a practitioner perspective, would have an empirical focus on the experiences and practices of those practising evaluation to inform, develop, and improve evaluation theory (Rog, 2015). This means that research on evaluation would accept that the problems we encounter in practice are ‘indeterminate’ and messy and not

always ideal for the application of evaluation theory, methods, approaches (Schwandt, 2014; Schön, 1991). As such, he argues that, *"the knowledge inherent in practice is to be understood as artful (skilful) doing … it is characterised by contextuality, acting that is continuous with knowing, and interactivity and that, taken together these elements comprise an account of the kind of practical judgment required of a professional who works in an environment characterised by complexity, indeterminacy, and the necessity to act on the situation at hand"* (Schwandt, 2014, p.233). Schwandt (2003) calls for research on practice more generally, and thus research on evaluation practice, to 'return to the study of the rough ground' and to take evaluation back to the 'real world' because in practice, evaluation situations are rarely perfectly suitable for following a particular theory to the letter (Smith, 1993). As such, the empirical focus of research on evaluation, using this perspective, would prioritise the practitioner perspective to help understand evaluation theory in practice.

Overall, when considering evaluation theory in evaluation practice, making the distinction between how we view how evaluation is being practised makes a difference to the empirical focus of research on programme evaluation. The instrumental rationality perspective takes a top-down approach to track the application of the principles or standards of an evaluation theory, method, or approach, whereas what I now refer to as the practitioner perspective places more emphasis on the language, behaviour, judgement, and perceptions of those practising evaluation to better understand 'the practical' of real-world evaluation practice.

2.3.4 Why is programme evaluation like this: the role of context

In conceptualising evaluation practice, a final matter to consider concerns the factors that shape why evaluation practice is the way that it is. Certainly, the purpose of evaluation and the questions evaluation activities seek to address are important drivers of how evaluation is conducted (Mark, Henry and Julnes,

2000). The ability of those practising evaluation to identify evaluation questions and utilise those to define evaluation activities which are feasible and useful has been discussed in the previous sections on ‘who is practising evaluation’ and ‘what is being practised’. However, there are other wider sociological and organisational contextual factors which can influence programme evaluation practice.

Understanding of ‘context’, as that which provides a fuller understanding about a given unit of analysis or target of inquiry (Vo and Christie, 2015), is an important factor to consider in research on programme evaluation. In the case of evaluation, context refers to the “landscape within which the evaluand, relevant stakeholders, and evaluator operate”. The landscape can include users of the evaluation, organisational, social, and political climates, values and interests, language and culture, and the way that all these factors interact (Vo and Christie, 2015; Rog, 2015). Understanding of context therefore points to important influences on evaluation processes and the resulting consequences of evaluation activities.

For example, evaluation practice can be perceived to be influenced by both organisational and social forces over time. Dahler-Larsen has written extensively on these ideas arguing that the conceptualisation of evaluation practice is associated with socio-historical expectations of evaluation practice (2016). He draws from both organisational and sociological theory to help explain how evaluation evolves and is practised as it is today (Dahler-Larsen, 2012, 2016) drawing heavily from the ideas of the ‘institutionalised

organisation'[12] and the 'audit society'[13] to explain that evaluation is in social demand due to organisations' desire to seek legitimacy, even when there is little learning or use from evaluation instrumentally within the organisation. The ideas behind the institutionalised organisation help to explain a lack of coherency in evaluation processes and why organisations might hesitate to use evaluation results. Dahler-Larsen states that this perspective explains evaluation practice in a more realistic way as opposed to other models of the organisation e.g. the simplistic rational model and the idealised learning model.

Dahler-Larsen draws also on sociological theory to understand better the social environments in which those practising evaluation in organisations are set. Sociologically, Dahler-Larsen points to the concept of the 'audit society' (Power, 1997) where performance-based accountability and monitoring is valued most and is embedded in reporting protocols by those contracting or funding the organisations that implement programmes, services, or interventions. On a practical level, evaluation practice is, according to this thinking, dependent upon the activities, institutions, politics, norms, values, and experiences existing within organisations and society (Dahler-Larsen, 2012). Dahler-Larsen's work serves to highlight the wider influences that shape evaluation practice across society.

[12] Evaluation is consistent with the "normative and cognitive [e.g. evaluation as an organisational procedure for control and management] pillars of institutional life in modern society. Increasingly, it is also mandated by the regulative pillar. This happens when some evaluation system is required by law, when evaluation is built into administrative policies, and when foundations support grantees only on the condition of subsequent evaluation" (Dahler-Larsen, 2012, p.68).

[13] The 'audit society' sees evaluation as a consequence of a society where rationality (i.e. predictability, objectivism, and procedure) is important and evaluation is associated with security and safety where evaluation procedures help to manage risk. The term 'evaluation machines' signifies evaluation in the audit society as value-free (Dahler-Larsen, 2012).

An acknowledgement of context therefore helps us to understand that evaluation practice is often not a rational, objective activity, rather that it is shaped by broader social forces, but also that the immediate organisational context of the evaluation impacts how evaluation is carried out. Acknowledging the role of context means that research on evaluation can better account for the socio-political as well as the external and internal organisational landscape which influence evaluation practice (Schwandt, 2003). In this respect, Schwandt states that we should accept that evaluation practice "is always carried out on a rough ground of paradox and contingency, ambiguity and fragmentation [that] what is really going on [in practice] is nothing but life forever unfolding on the rough ground and, thus, indelibly marked by distinctive tensions, contradictions, paradoxes and dilemmas that affect our understandings of self, world and other, and consequently, our practices" (Schwandt, 2003, p.361).

2.3.4.1 Implications for research on programme evaluation

Research on evaluation practice can vary in the extent to which it accounts for the context of evaluation practice. Research on evaluation which focuses on how an evaluation serves its intended purpose or answers specific evaluation questions might be less inclined to address the distinct aspects of the organisational and social context that shapes evaluation practice. However, such contextual information can help us to better understand why and under what social and organisational conditions specific evaluation theories, methods, or approaches may be advantageous or problematic to implement.

2.4 A conceptual framework of evaluation practice

I have used the first half of this chapter to lay out the various ways in which evaluation practice can be conceptualised and how each conceptualisation impacts the focus and methodological approach of research on evaluation

practice. What is clear from the preceding section of the review is that each conceptualisation of evaluation practice relating to the questions of who is practising programme evaluation, what is being practised, how programme evaluation is being practised, and the role of context, has distinct value for research on programme evaluation. As such, it is not that one conceptualisation reflects practice 'better' than any other, rather it is that different conceptualisations offer different ways of thinking about research on programme evaluation, and subsequently offers different ways of identifying research gaps, of conducting research, and of generating different forms of learning about programme evaluation practice. To guide the second half of this literature review, I have summarised the various approaches to conceptualising evaluation practice in Table 2. I use this framework to structure, inform and map the review of literature on the use of programme theory in programme evaluation practice. It should be noted that in Table 2, there are no distinct links between the items in each column, rather it simply provides the conceptualisations discussed in the preceding paragraphs across each row.

Defining questions about programme evaluation practice	**Ways to conceptualise programme evaluation practice**	
Who?	Accidental evaluator	Trained evaluator
What?	Formal programme evaluation activities	Evaluative thinking
How?	Instrumental rationality	Practitioner perspective
Why?	Role of context	

Table 2 Conceptualising evaluation practice - a summary

2.5 Conceptualising the use of programme theory in evaluation practice

I now move to applying that which has been discussed in the first half of this chapter to the specific case of the use of programme theory in programme

evaluation practice. I utilise the same questions that I addressed previously to structure this review of the literature. I therefore address the questions of who is using programme theory in programme evaluation practice; in what programme evaluation practices is programme theory used; how is the use of programme theory in evaluation practice conceptualised; and the contextual influences on the use of programme theory in programme evaluation practice. In this half of the literature review, I draw from both theoretical and conceptual literature as well as empirical studies on the use of programme theory in programme evaluation practice.

2.5.1 Who is using programme theory in evaluation practice?

The question of who is using programme theory in programme evaluation practice is not always made explicit in the literature. However, there are some indications about this. For example, Coryn and colleagues (2011) find that applications of 'theory-based evaluation' are often carried out within the context of large research-focussed projects, rather than in smaller settings, which often implies that the primary evaluation practitioner is formally trained in evaluation or social science methodologies. An example of this is the study by Deane & Harré who report on an extensive programme evaluation which aimed to test the programme theory using empirical data collection (2014). The programme evaluated in this case is a youth development programme involving adventure, service-learning, and mentoring components. In developing programme theory, the authors triangulate multiple sources of information including the perspectives of programme stakeholders (programme management and programme participants). This programme theory was then tested against relevant academic theories, empirical evidence, and expert opinions. The authors, from the University of Auckland, however report that it was they who directed and conducted most of the empirical work with respect to the evaluation. Moreover, there is little space given in the article to noting the utility of this approach to evaluation for evaluation stakeholders, other than a brief mention that the use of programme theory helped to identify some areas for programme improvement.

Certainly the involvement of *evaluation stakeholders* is emphasised in the development and use of programme theory in evaluation practice more generally (Huebner, 2000) to increase buy-in to the evaluation approach (Sullivan, Barnes and Matka, 2002; Sullivan and Stewart, 2006) and to increase the utility of the evaluation findings more broadly (Friedman, 2001; Patton, 2012). Academics in this space note the fine balance between the ability of programme theory to account for the dynamic and complex nature of programmes or interventions (Mills, Lawton and Sheard, 2019) while remaining accessible and useful for such stakeholders (Davies, 2018). However, overall in this literature, the main 'users' of programme theory, as in those who develop and use programme theory in the context of evaluation, appear to be predominantly either social science researchers or professional evaluators, often working *with* stakeholders to conduct programme evaluation, rather than say the use of programme theory by 'accidental evaluators'.

Another reason for this observation might be the fact that those reporting on evaluation for academic publication are more likely to be academics, social science researchers, or professional evaluators, rather than say an 'accidental evaluator' for whom evaluation is but one part of their job. In such cases, publishing their evaluations in an academic outlet is not necessarily a priority. This means however, that the programme theory literature does not reflect the full spectrum of those who practise evaluation, i.e. those who are classed as 'accidental evaluators'. Our understanding of the use of programme theory may therefore be somewhat limited to the perspectives and preferences of trained evaluators.

Nevertheless, the *potential* utility of using programme theory in settings which are not large well-funded research projects is noted. Some argue that theory-based approaches for small practitioner-based programmes funded at low-

levels can still have relevance and advantages (Birckmayer and Weiss, 2000). The authors here point to potential ways in which theory-based approaches can be implemented in such settings including: using programme theory even if it is not yet fully developed; considering programme theory development as a stage in evaluation; including a process evaluation; and using the information that a theory-based evaluation can provide (Birckmayer and Weiss, 2000). This stands in contrast to the perspective of some advocates of programme theory approaches who argue that the principles of a given approach should not be applied 'slavishly' (Marchal et al., 2012) i.e. evaluations should adhere to the key principles of the specific approach, and not apply it in a piecemeal way. Nevertheless, the perspective of small practitioner-based programmes is still lacking in the empirical literature.

2.5.2 What programme evaluation practices is programme theory used within?

The use of programme theory in formal (systematic) evaluation activities

It is quite clear from the literature than the predominant use of programme theory has been conceptualised from the perspective of the benefits it affords to the conduct of formal (systematic) evaluation activities, particularly from a methodological perspective. This is a likely consequence of the original purpose of developing and using programme theory to open up the 'black-box' of programmes to understand how and why outcomes change, or not. Several of the original intentions of the use of programme theory approaches include one of, or a combination of, the following methodological benefits of the use of programme theory:

1. to increase the *internal validity* of evaluations (Palumbo and Oliverio, 1989) through understanding how and why programmes achieve desired change or not (Weiss, 1997b) rather than simply stating whether or not an outcome was changed;

2. to increase the external validity of cumulative evaluation knowledge to enable the *replication* of successful programmes (Hacsi, 2000) through generating understanding about the key processes and mechanisms through which change occurs;
3. and to build theory about successful strategies to tackle problems facing society (Pawson and Tilley, 1997; Astbury and Leeuw, 2010) through cumulative knowledge building.

As a result of the original intentions for the use of programme theory, the ways in which programme theory can be used for such purposes are clear from the literature. This is a likely consequence of the emergence of the development of various methodological approaches to implementing programme theory in formal evaluation activities. Examples of such approaches are: theory-driven evaluation (Chen and Rossi, 1980), theory-based evaluation (Weiss, 1995), realist evaluation (Pawson and Tilley, 1997), contribution analysis (Mayne, 1999), and logic analysis (Brousselle and Champagne, 2011): I provide a more in-depth summary of the approaches in the table given in Appendix 1. Each of these approaches has a set of guidelines, or principles, for developing and using programme theory in the case of formal (systematic) evaluation activities. In most cases, the focus is on the use of programme theory to deal with complicated and complex aspects of programmes and interventions (Weiss, 1995, 1997a; Stame, 2004; Rogers, 2008; Dalkin et al., 2018; Mills, Lawton and Sheard, 2019), especially for understanding how change occurs and if not, at which point in the process it breaks down (Weiss, 1997a), i.e. to 'test' the programme theory. Testing programme theory in this way can help to identify whether the programme was implemented badly or whether the programme as a whole is not fit for purpose to address the changes in outcomes targeted (also known as implementation failure and programme failure respectively) (Lipsey and Pollard, 1989).

However, the practice of programme theory has evolved over the last 30 years to include support for its use in other formal (systematic) evaluation activities (Funnell and Rogers, 2011) such as planning/designing programmes, services, and interventions (Julian, Jones and Deyo, 1995; Julian, 1997; Keller and Bauerle, 2009), the evaluation of processes and implementation of programmes (Peyton and Scicchitano, 2017), to support organisational learning about the programme (Holma and Kontinen, 2011), outcome and impact evaluations (Cooksy, Gill and Kelly, 2001; Wimbush, Montague and Mulherin, 2012), and performance monitoring/measurement at the programme and system level (McLaughlin and Jordan, 1999; Mayne, 1999; Funnell, 2000; Millar, Simeone and Carnevale, 2001; Reed and Jordan, 2007). In the case of performance monitoring/measurement, the role of tools such as the logic model have been emphasised in the literature (McLaughlin and Jordan, 1999; Taylor-Powell, Jones and Henert, 2003; Millar, Simeone and Carnevale, 2001). Logic models are useful in this sense as they provide a means to mapping programme inputs alongside programme activities, outputs, and outcomes with the goal of facilitating accurate monitoring and evaluation (Frechtling, 2007; Taylor-Powell, Jones and Henert, 2003) to allow more effective identification of indicators of measurement of performance and success of programmes.

Furthermore, empirical studies of the application of programme theory have also focussed their research questions on the use of programme theory in formal (systematic) evaluation activities (Birckmayer and Weiss, 2000; Coryn et al., 2011). As an example, one of the primary research questions addressed in a systematic review of practice (Coryn et al., 2011) was “for what purposes (e.g., formative, summative, and knowledge generation) are theory-driven evaluations conducted?”. To summarise, considering these approaches together, the use of programme theory generally involves making explicit the underlying assumptions about how and why programmes are expected to work and using this information to guide formal evaluation activities in some way (Rogers, Petrosino, Huebner and Hacsi, 2000).

However, according to empirical reviews of practice reported in the literature (Birckmayer and Weiss, 2000; Coryn et al., 2011), it was claimed that, even in cases where programme theory had claimed to be used in formal evaluation activities, the added value of doing so is not often clear. It could be that empirical studies have simply not reported on the added value of using the programme theory in formal (systematic) evaluation activities, even when there was value. It could be that those writing up the evaluation may have not considered it important to mention the value of programme theory because they were primarily concerned with writing up the findings of the evaluation rather than focussing on the role of programme theory. It seems more likely that the authors preferred to disseminate the evaluation methods and findings, particularly when publishing in fields outside of the evaluation field, which is often the case as programme evaluations are often discipline-specific. It follows that what we see, published in the literature, is often more focused on the *use* of programme theory in formal (systematic) evaluation activities, rather than a focus on *how* the programme theory helped evaluators to think evaluatively for example.

The use of programme theory in evaluative thinking

In terms of evaluative thinking, i.e. the thinking skills required to define, frame, design, and carry out evaluation, there is less explicit focus on conceptualising the use of programme theory in practice in this way compared with the conceptualisation of the use of programme theory in formal (systematic) programme evaluation activities. By explicit focus, I mean that it is not clear *how* the use of programme theory facilitates evaluative thinking in the literature. However, it seems to be implied that the use of programme theory can facilitate evaluative thinking through the use of programme theory to guide evaluation in defining and prioritising evaluation questions, data collection, and data analysis (Birckmayer and Weiss, 2000; Donaldson, 2007). That being said, an empirical review of the use of theory-based approaches in evaluation

highlights that even in cases where the use of such approaches is reported, the value of doing so was not always clear (Birckmayer and Weiss, 2000; Coryn et al., 2011). This is due, at least in part, to a lack of understanding about *how* the use of programme theory can help to define, frame, and conduct formal evaluation activities through facilitating evaluative thinking. Some examples from the literature deal with this question of how the use of programme theory can facilitate evaluative thinking, in many cases implicitly. I discuss several of these examples below.

An example of the consideration of the role of programme theory in facilitating evaluative thinking is in case studies of the application of contribution analysis in policy evaluation presented by Wimbush and colleagues (2012). Whilst focused on the evaluation of policy (a larger scale than might be expected in programme evaluation in small TSOs), the examples presented in this study help to conceptualise the role of programme theory (at least through contribution analysis) for evaluative thinking. The findings state that, as a participatory and structured process, the use of contribution analysis provides a *lens* and a *common language* to strengthen collaborative capacity to plan evaluation and strengthen outcomes thinking and reporting, primarily through the participatory process to developing the theory-of-change (and programme theory). In this sense, the authors emphasise the value of the process of developing programme theory in developing a shared understanding of the theory and using this shared understanding to strengthen evaluation planning. However, this example is based on an evaluation within the much broader context of policy evaluation and using programme theory, as but one part of a theory-of-change methodological approach.

Other examples, exploring the use of programme theory for evaluative thinking more implicitly also exist. Of note is the case of ‘evaluability assessment’. Whilst programme theory forms only one part of the evaluability assessment process, the goal of evaluability assessment is to ensure that programmes are

properly defined and implemented before conducting systematic evaluation activities (Trevisan, 2007). As such, it is now recognised as a way to balance increasing demands for evaluation with limited resources, through facilitating better preparation and readiness for evaluation (Craig and Campbell, 2015). In terms of evaluative thinking, the approach overall aims to facilitate the framing, design and conduct of systematic evaluation activities that are feasible and useful to programme stakeholders, ultimately ensuring that formal (systematic) evaluation activities are only conducted when the programme is 'ready' to be evaluated.

Originally, six key steps for undertaking evaluability assessment were outlined. These include: involving intended users of evaluation information, clarifying the intended programme, exploring the implementation of the programme, reaching agreement on any changes in activities/goals, exploring different evaluation designs, and agreeing on evaluation priorities as well as intended uses of the evaluation (Leviton et al., 2010, p.217). A key part of this process of evaluability assessment involves developing a programme model, or theory, or logic model, or theory-of-change (Wholey, 1987; Leviton et al., 2010; Craig and Campbell, 2015). Despite the concern over declining evaluability assessments in the published literature, a review conducted by Trevisan locates over 20 examples of evaluability assessment in the literature from 1986 to 2006, with a wide variety of disciplines represented. Moreover, evaluability assessment has been studied in more detail in specific disciplines e.g. the impact of evaluability assessment for public health programmes, policies and practice (Leviton et al., 2010).

Other examples of a more implicit consideration of the use of programme theory for evaluative thinking exist. Based on a review of evaluation documents, from evaluations conducted by four different evaluation teams, Huebner comments on the value of using programme theory evaluation in schools (2000). She states that in the cases reviewed, the development of

programme theory (in a logic model format) firstly encourages *cooperation* and *buy-in* from staff through *participation* in the programme theory building process. She reports that this encouraged staff to be more receptive and willing to participate in evaluation. She reports that the development of programme theory encouraged *reflective practice*, particularly reflection-on-action (Schön, 1991), where teachers were able to use the process of developing programme theory to reflect on the ways in which their work contributed to improved learning outcomes for pupils in the classroom, which can then be reflected in programme evaluation. Whilst this example is not explicit about the role of programme theory for evaluative thinking, it does highlight that the process of developing programme theory may have benefits for supporting evaluative thinking in allowing practitioners to reflect on their work and use this to inform formal (systematic) evaluation activities.

In another paper, Friedman highlights the potential role that can be played by developing programme theory in developing *consensus* amongst staff through confronting 'designed blindness'. 'Designed blindness' essentially reflects the gaps between stakeholders' intended programme theory (espoused theory) and actual implementation (theory-in-use) (Argyris and Schön, 1996). Friedman states that using programme theory to overcome designed blindness can improve therefore improve the quality of information (i.e. programme theory) upon which formal evaluation activities are conducted (2001). Surfacing programme theory in this example leads to staff members being more aware of their implicit assumptions about how the programme was working. In confronting designed blindness, Friedman states that staff can develop consensus about programme implementation and identify where there might be barriers to effective programme implementation.

Often when considered implicitly, the role of programme theory in evaluative thinking is not the primary empirical focus of the article. A lack of proper and explicit conceptualisation of the role of programme theory in evaluative thinking

makes it difficult to draw learning from across studies that consider the role of programme theory in this respect. However, from the empirical studies that do exist, we can see the potential value in the role of developing programme theory in supporting evaluative thinking, primarily in employing a participatory process to develop a common or shared understanding of the programme theory through the opportunity to reflect-on-action. In the examples given above, the common or shared understanding can encourage buy-in and willingness from staff to participate in evaluation, can provide better quality information for planning evaluation activities, and can strengthen the capacity to develop outcomes and plan evaluation. However, this conceptualisation of the role of programme theory is partly based on studies which use secondary reports of evaluations as a data source; this may not accurately report on the role of programme theory in evaluative thinking. Some studies are also of nation-wide policy interventions, and as such do not reflect work at the programme level. Overall, the literature on the use of programme theory for evaluative thinking focusses more systematically on *what* it helped organisations to do, rather than *how* it helped them to do it.

2.5.3 How is the use of programme theory in evaluation conceptualised?

The methods used to conduct research on the use of programme theory in practice predominantly reflect the instrumental rationality perspective on how evaluation is practised. Much of the empirical literature which would be classed as 'research on evaluation' uses a literature review methodology including both structured and systematic literature review methods. The focus of empirical literature reviews of how programme theory is used in evaluation practice adopt an 'instrumental rationality' perspective because the focus is on how theory was applied in practice. Many empirical articles do this by tracking the application of methodological principles and concepts through reviews of studies that have applied programme theory approaches (Coryn et al., 2011; Marchal et al., 2012; Vogel, 2012; Lacouture, Breton, Guichard and Ridde, 2015; Dalkin et al., 2015). One of the most substantial reviews conducted is by Chris Coryn and colleagues where they define the use of programme theory

to be “any evaluation strategy or approach that explicitly integrates and uses stakeholder, social science, some combination of, or other types of theories in conceptualising, designing, conducting, interpreting, and applying an evaluation” (Coryn et al., 2011, p.201). The review examines 45 cases of ‘theory-driven evaluations’ in total to identify how closely the practices adhere to the core principles of theory-driven evaluation. Generally, these studies give us an understanding of the extent to which the principles of specific programme theory approaches are applied, or reported.

Similar studies provide the reader with good understanding of how methodological principles and concepts are applied, or are reported upon, in certain settings e.g. health (Marchal et al., 2012; Breuer, Lee, De Silva and Lund, 2016), with respect to specific approaches to using programme theory e.g. realist evaluation (Lacouture et al., 2015). With such approaches to understanding practice, the extent to which methodological principles are adhered to acts as the normative benchmark for success of the use of programme theory. This perspective on how programme theory is used in evaluation sees practice as more rational and analytic (Schwandt, 1989), where the goal is the application of theory in practice. This dominant perspective lacks consideration from a practitioner perspective.

Moreover, by assuming that full methodological implementation of an approach is the benchmark for successful use of programme theory, I argue that we limit our understanding of the potential role that can be played by programme theory in the context of evaluation. For example, Patricia Rogers cites Weick’s assertion that even if the model (programme theory) is not entirely correct that it can still provide a useful heuristic tool for purposeful action because “once people begin to act … this helps them to discover what is occurring, what needs to be explained, and what should be done next” (Weick, 1995, pp.54–55). In Weick’s words, having some form of model at least brings order to the world which might then stimulate action (Weick, 1995).

Rogers and Weick therefore emphasise the question as to what happens when programme theory is used under less strict conditions than those implied by more structured approaches to its implementation, e.g. through the use of methodological guidelines, such as in the case of logic analysis or realist evaluation.

Methodologically, the use of structured and systematic literature reviews (Coryn et al., 2011; Trevisan, 2007; Marchal et al., 2012; Rogers et al., 2000) provide an depth understanding of the scope and findings of published peer-reviewed applications of the use of programme theory. However, the extent to which the search strategy is detailed varies and thus the reliability of the articles reviewed can be questioned. Such search strategies can introduce bias to certain types of evaluation, i.e. those likely to be published in peer reviewed journals, and this contributes to our limited view on the role of programme theory in evaluation practice.

A further, more practical, challenge with using structured and systematic review methodologies in research on evaluation is that, as with any professional and practice-oriented discipline such as programme evaluation, there is a substantial body of grey literature where programme evaluations might be published: such evaluations are often not included within the structured literature search strategy. In many cases it is not feasible to search the entirety of the grey literature other than in those cases where the review is focussed on a specific intervention or programme type for example: in such cases, relevant sources of grey literature can be more easily identified. However, for studies focusing on the approach (programme theory) more widely, it is much more difficult to bound a literature search strategy in the grey literature.

On the other hand, we see many anecdotal case study reviews of the practice of developing and using programme theory (Birckmayer and Weiss, 2000; Sullivan, Barnes and Matka, 2002). In these cases, authors have either chosen several evaluations to review with no search strategy outlined, or have chosen to review the application of a particular approach across a programme setting (e.g. Health Action Zones). This methodological approach, albeit less systematic, offers in-depth learning of the application of programme theory such that findings reflect what was learned, what the benefits were, and challenges to its implementation (Birckmayer and Weiss, 2000). For example, in the review by Birckmayer and Weiss, the reported benefits of the use of programme theory are in terms of informing programme improvement through identifying critical programme components; for supporting knowledge generation about social change; and in terms of evaluation planning in identifying clear guidelines for data collection and analysis (2000). The challenges in implementing programme theory related to how well the programme theory is defined, how well programme activities reflect the assumptions about how the programme works, and resources (money and time). Whilst informative about the potentially valuable role played by programme theory, this descriptive reporting in an often indirect way, i.e. through reviewing secondary evaluation reports, means that such articles often lack a clear framework from which the reader can draw learning from across several studies.

Overall, I conclude that structured or systematic literature reviews of the use of programme theory from the perspective of methodological implementation, such as those discussed in the preceding paragraphs, variably report the application of methodological principles. Other than analysing implementation of the approach according to the principles of, or concepts within, that approach (Coryn et al., 2011; Trevisan, 2007; Marchal et al., 2012), authors generally tend to describe anecdotally the learning derived from their own perspective, often within the context of the wider evaluation methods and findings. Whilst this is certainly informative, the lack of a clear framework,

focussing on the role of programme theory, makes it difficult to draw learning out of these studies, across programme settings, and across different programme theory approaches. Moreover, the lack of framework questions whether specific challenges or benefits in the use of programme theory have been omitted. When the primary focus is not on the role of programme theory, then what we can know about the role of programme theory is dependent on what is written up, reported upon, or considered important by authors, rather than that which is guided by a clear framework.

2.5.4 Evaluation practice context and the use of programme theory

Many applications of programme theory, discussed in the previous section, are from the perspective of large research projects (Coryn et al., 2011). As a result, the purpose of the use of programme theory is often to develop knowledge about some programme or intervention type. The original purpose served by the use of programme theory, in the context of evaluation, was opening up the 'black box' of programmes (Chen and Rossi, 1980, 1987). Consequently, from a social science perspective, such approaches to using programme theory have proved popular tools for developing knowledge about how and why programmes work by exploring the "space between the actual input and the expected output of a programme" (Stame, 2004, p.58). Whilst this offers the opportunity to increase the internal and external validity of programme evaluations, and generate learning about specific programmes and their context, there is less consideration of the context in which programme theory in being used i.e. the context of programme evaluation practice.

From the perspective of evaluation practitioners, there is often little consideration of contextual influences on the use of programme theory. Certainly, applications often report that utilising programme theory in formal (systematic) evaluation activities is time, resource, and expertise intensive (Weiss, 1997a; Mayne, 1999; Rogers et al., 2000; Birckmayer and Weiss, 2000; Blamey and Mackenzie, 2007) and that these internal resource-based

factors may inhibit the use of programme theory. This tells us less about the external context affecting programme evaluation practice.

An example, which explicitly considers the role of the wider external context, with respect to Theory of Change evaluation, is a study of its use in Health Action Zone (HAZ) evaluations. Sullivan and Stewart (2006), for example, note, in their case study of the application of theory of change evaluation to national Health Action Zone (HAZ) evaluations in the UK, that context plays an important role in mediating the utility of the approach, not just the evaluation resources or evaluation expertise. They note that theory of change evaluation is perhaps not universally transferable to the UK for two socio-political reasons. First, evaluation is often considered as a post-hoc activity i.e. there is often a lack of attention given to evaluation in earlier stages of the programme life cycle e.g. in programme design. Second, they find that the political structure (i.e. the centrality of government) is a constraining factor to the development and use of theory of change evaluation for policy analysis given that policies are centrally defined and the theories of change developed centrally are not reflective of the local circumstances in which the theories of change are to be implemented (Sullivan and Stewart, 2006). These findings stand in contrast to the fact that UK government evaluation guidance advocates the use of programme theory[14] early in the programme lifecycle. These contextual challenges mean that in this context, the authors argue, that the utility of theory of change evaluation cannot meet the full potential originally intended by the Aspen Institute. On the other hand, many conceptualisations of the use of programme theory recognise its value in its ability to facilitate reporting in terms of meeting accountability requirements, particularly with the use of logic models (Hernandez, 2000; Hatry et al., 1996). In this respect, accountability is

[14] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

a strong contextual narrative with respect to programme evaluation practice in the public and third sectors.

There could be value in better understanding the contextual influences on the use of programme theory in practice in terms of considering the social and organisational contexts within which those practising evaluation are situated and how these contextual factors might impact upon the choice to use programme theory, and the value of its use. Making these contextual factors a focus in research on evaluation can illuminate in more detail how and why the use of programme theory may be advantageous or not in specific evaluation contexts, much like in the case presented by Sullivan and Stewart (2006). Understanding of contextual factors affecting the use of programme theory may, in some cases, address some of the challenges highlighted in the literature where the added value of using programme theory was not clear in many instances (Birckmayer and Weiss, 2000; Coryn et al., 2011)

2.5.5 Conceptualising the use of programme theory in the context of evaluation

To summarise the latter half of this chapter, I have explored the literature that conceptualises the use of programme theory in programme evaluation practice. This is from the perspective of both theoretical and conceptual literature, as well as empirical studies analysing the use of programme theory in programme evaluation practice. I address several important questions with respect to conceptualising programme evaluation practice that affect what aspects of evaluation practice are studied, and subsequently, how we can go about conducting research on programme evaluation practice. Specifically, I address the questions of: who is using programme theory in programme evaluation practice; within what programme evaluation practices is programme theory used; how is the use of programme theory in programme evaluation practice conceptualised in empirical studies; and the contextual influences on the use of programme theory in programme evaluation practice.

In addressing these four questions, I can characterise the literature on the use of programme theory in practice as follows:

1. This literature is primarily focussed on applications of programme theory in large-research projects where we can assume that those seeking to publish the findings of such projects are those either trained in evaluation methods, or those with training in social science methodology. Whilst it is clear that stakeholder involvement is important in developing and using programme theory, the voice of evaluation stakeholders or accidental evaluators are not prioritised in this body of literature.
2. In terms of what programme evaluation activities programme theory is used for, there is much focus on its use in formal (systematic) evaluation activities with only some explicit and implicit consideration of its use in terms of facilitating evaluative thinking. However, most of this literature focuses on what the development of programme theory can help evaluators to do, rather than how it helps them do it.
3. In terms of how programme theory is practised, those studying its use adopt primarily an instrumental rationality perspective that seeks to use traditional social science methods, such as structured and systematic literature reviews, to track the application of methodological principles. Moreover, descriptive and anecdotal accounts of evaluations which use programme theory, whilst numerous, lack a consistent reporting framework and as such, it is difficult to draw out learning about the role of programme theory across studies.
4. I find that given that the literature broadly consists of larger research project applications of programme theory, the focus is on generating learning about the programme and not about the role of programme theory in the context of programme evaluation practice. Therefore there is little consideration of the social, political, and organisational contexts influencing the use of programme theory in practice.

Using the conceptual framework, I first introduced in Table 2, earlier in this chapter, I have summarised the latter half of this chapter using a 'traffic light' system in Table 3. The green shaded boxes reflect the predominant perspective in the literature on the use of programme theory in programme evaluation practice. The red shaded boxes highlight where the conceptualisations of programme evaluation practice show little focus. The amber box reflects that some consideration of the use of programme theory is reflected.

Defining questions about programme evaluation practice	**Ways to conceptualise programme evaluation practice**	
Who?	Accidental evaluator	Trained evaluator
What?	Formal programme evaluation activities	Evaluative thinking
How?	Instrumental rationality	Practitioner perspective
Why?	Role of contextual influences	

Table 3 Conceptualising the use of programme theory in evaluation practice - a summary

2.6 Identifying the research problem

In the introductory chapter of this thesis, I outlined that I am interested in the use of programme theory in the context of small TSOs. In comparing what we know about evaluation practice in such organisations (Chapter 1) with that which I have just discussed on the use of programme theory in programme evaluation practice (summarised in Table 3), there are a number of challenges which the literature does not address.

1. A large proportion of those practising evaluation in small TSOs are classed as 'accidental evaluators' i.e. those without formal evaluation training. The literature discussed in this chapter does not pay much

focus to the use of programme theory by 'accidental evaluators'. Whilst it heavily emphasises stakeholder involvement in the development and use of programme theory, the majority of reporting is from the perspective of the trained evaluator, and is inconsistent and does not necessarily prioritise the voices of the stakeholders involved, primarily reporting instead on the findings of the evaluation itself. Moreover, with a focus on the extent of application of methodological principles, the literature discussed takes little account of the implicit theories guiding evaluation practice with respect to the use of programme theory. In the case of small TSOs I argue that both capacity concerns of accidental evaluators, and their implicit theories of evaluation practice, are important factors to consider.

2. The social problems which small TSOs are increasingly charged with tackling through their programmes and services are challenging to evaluate. This is because such social problems are often caused by and manifested in multiple other problems. As a result, it can be difficult to isolate the difference or changes that a programme or service has contributed to because the social issues are so contextually embedded. Having the capacity to evaluate such programmes in terms of both knowledge and skills, and the ability to 'think evaluatively', is important in these settings where those practising evaluation are based within the organisation with little or no formal evaluation training. The literature discussed in this chapter pays little explicit attention to the role of programme theory in facilitating evaluative thinking, i.e. how it helps evaluators to define, frame, and carry out formal (systematic) evaluation activities. It may be that in cases, such as small TSOs, where substantive knowledge of the field and programme is likely high, but evaluation expertise low, that supporting evaluative thinking skills is important in order to link expertise in the subject area with the conduct of formal evaluation activities.
3. Small TSOs are operating within dynamic internal and external organisational environments which present a number of contextual

challenges in terms of evaluation. These challenges include, but are not limited to, increasing accountability requirements, complex funding mechanisms, and the need to ensure long-term sustainability through organisational learning, amongst others. This means that the practice of evaluation needs to serve several purposes as a result of this contextual environment to help small TSOs to remain accountable whilst maintaining both financial sustainability and effective service provision. The literature discussed in this chapter takes little account of how these kinds of contextual influences might influence the use of programme theory in programme evaluation practice. However, in contexts, such as small TSOs, that often conduct evaluation outside of social science or academic research contexts, their organisational, social, and political contexts have a large influence in shaping and defining their practices both in terms of evaluation, and their day-to-day service provision practices.

2.6.1 Specifying research questions

It follows that the understanding of the use of programme theory in programme evaluation practice is not consistent with the image of evaluation practice in small TSOs, which I discussed in Chapter 1. Therefore, I propose to address two research questions which aim to explore and develop the understanding of the current and potential role of programme theory in small TSOs, with the overall aim of better conceptualising the use of programme theory in small TSOs and closing the gaps between theory and practice. The two research questions addressed in this thesis are:

RQ1: What is the current practice with respect to the use of programme theory in programme evaluation practice in small TSOs?

a. What is the context of programme evaluation practice in small TSOs with respect to the capacity to do evaluation, evaluation activities, and the use of evaluation?

b. In what forms, does the use of programme theory manifest itself within the context of programme evaluation practice in small TSOs?

RQ2: How can the development of programme theory improve future programme evaluation practices within small TSOs (particularly with respect to facilitating evaluative thinking)?

In the following chapter, I define and justify the methodological framework I used to address these research questions.

CHAPTER 3 METHODOLOGICAL FRAMEWORK

3.1 Introduction

In Chapter 1, I highlighted the importance of conducting research on evaluation as a means to better aligning theory and practice in the programme evaluation field. I also introduced my interest in both programme evaluation practice in small TSOs, and the use of programme theory in programme evaluation practice. In Chapter 2, I argue that how evaluation practice is defined will shape that which research on evaluation focuses on, how research on evaluation is conducted, and thus the findings from research on evaluation that emerge. I outline the differing conceptualisations of evaluation practice according to the questions of: who is practising evaluation; what is being practised; how is programme evaluation practice is conceptualised; and why evaluation being practised in the way that it is.

In the case of the use of programme theory in programme evaluation practice, I argue that its current conceptualisation does not align with programme evaluation practice in the context of small TSOs. I argue that the literature on the use of programme theory in evaluation practice is predominantly, although not exclusively, focused on the use of programme theory by trained evaluators, or social scientists, and in formal (systematic) evaluation activities, rather than in evaluative thinking. Moreover, research on the use of programme theory in practice is defined primarily as an application of methodological guidelines or principles, rather than the use of programme theory as understood from a practitioner perspective. Contextual influences on the use of programme theory in programme evaluation practice are also given little emphasis in the literature. Rather, the literature focuses on exploring the context of the programme itself as opposed to the context of carrying out the evaluation.

In Chapter 2, I therefore argue that the conceptualisation of programme evaluation practice with respect to the use of programme theory is not consistent with what we understand about programme evaluation practice in

small TSOs for a number of reasons, including that those practising evaluation are often ‘accidental evaluators’ and as such may not be aware of or use evaluation theory in their practice. This means that their practice is likely to be defined or influenced by other factors, including practitioners’ ‘implicit theories’ of evaluation, professional judgement, values, and experience, internal and external organisational environments, and importantly, the capacity to do evaluation.

I therefore identify the two research questions that the empirical component of this thesis addresses with the aim of better aligning theory and practice with respect to the use of programme theory in programme evaluation practice in small TSOs.

3.1.1 Research questions

RQ1: What is the current practice with respect to the use of programme theory in programme evaluation practice in small TSOs?

a. What is the context of programme evaluation practice in small TSOs with respect to the capacity to do evaluation, evaluation activities, and the use of evaluation?
b. In what forms, does the use of programme theory manifest itself within the context of programme evaluation practice in small TSOs?

RQ2: How can the development of programme theory improve future programme evaluation practices within small TSOs (particularly with respect to facilitating evaluative thinking)?

3.1.2 Chapter structure

This chapter defines and justifies the methodological framework underpinning the empirical component of this thesis. As such, this chapter is structured as follows. First, I outline the practitioner-oriented theoretical lens through which

I conduct this research. Second, I define the pragmatic approach to conducting research. Lastly, I define the multi-methodological framework used to address the two research questions. In this section, I detail the specific methods I use to conduct two empirical studies.

3.2 Theoretical lens - a practitioner-oriented perspective

In Chapter 2, I identify that the literature on the use of programme theory in programme evaluation practice is conceptualised from a perspective that primarily emphasises the role of the trained evaluator and reports on the methodological implementation of programme theory in formal (systematic) evaluation activities. I argue that this does not align with evaluation practice in small TSOs. It is not that this dominant perspective is not useful, rather it is not necessarily wholly reflective of programme evaluation practice in different contexts and settings.

Therefore, the focus of the empirical component of this research is programme evaluation practice from a perspective which values the 'implicit theories' and the resulting day-to-day practices of those practising evaluation in small TSOs. Implicit theories reflect how the perspectives of practitioners on evaluation guide their evaluation work (Christie, 2003b). This is important because given the lack of formal evaluation training, it is possible that evaluation theories may not guide programme evaluation practice at all. The implication of this, as I discussed in Chapter 2, is that evaluation practice is then driven by practitioners' 'implicit theories' of how to go about conducting evaluation including, but not limited to, practitioner habit, intuition, judgement, and values, organisational context, and the capacity to do evaluation (Dahler-Larsen, 2014).

Scholars who view evaluation practice from the practitioner perspective commonly draw from Donald Schön's *Theory of the Reflective Practitioner*

(Schön, 1991). Schön's thinking is particularly relevant here because he highlights the value of the practitioner perspective. He argues that professions are central to the functioning of society. However, he also argues that despite the centrality and importance of professions in society, for many reasons, including the abuse of autonomy of professionals, research has not prioritised the knowledge of professionals, and that the dominant epistemology in research has been the application of scientific theory and technique to understand practice (Schön, 1991). This dominant positivist epistemology of practice relates to the view of evaluation practice called 'instrumental rationality' (Dahler-Larsen, 2012).

Adopting the instrumental rationality perspective means that in the evaluation field, and social sciences more generally, we often see a division between theory and practice because the perspectives and methods we use to conduct research do not adequately account for the knowledge and perspectives of professionals. In the broader field of organisation and management research, this dominant perspective has also been criticised, not because it is invalid per se, rather because it omits the role of *managing*. In his work on the topic, Mintzberg argues, for example, that a predominant focus on developing new programmes to tackle social issues means the field of management has ignored the role of managers and implementers in ensuring such programmes are adequately implemented and evaluated through the process of managing (Mintzberg, 1973, 2009). The omission of a practitioner perspective in the literature on the use of programme theory in programme evaluation practice further highlights the challenges noted by Mintzberg: this literature presents a strong focus on the use of programme theory in terms of its benefits for generating knowledge about a specific programme, rather than its benefits for actually conducting the evaluation activities.

As a result of the lack of consideration from a practitioner perspective, the impact of research on evaluation has been questioned (Coryn et al., 2016). As

such, persistent gaps between theory and practice can often persist in the evaluation field. Adopting a more practitioner-oriented perspective would, according to the thinking outlined in the preceding paragraph, facilitate a more complementary dialogue between evaluation theory and practice in many cases (Schwandt, 2003).

From a reflective practitioner perspective, the value of placing practitioners, and their implicit theories of practice, as the focus of research on evaluation is similarly related to the thinking on *pragmatism* as a 'philosophy of practice' (Simpson, 2018). Pragmatism in this sense perceives the ideas of thinking and doing as often inseparable in practices which are continuously evolving (Mintzberg, 1973; Schön, 1991; Simpson, 2018) and as such, views practice as situated and social, and as shaping and being shaped by the lived experience of knowers (Simpson, 2018). In terms of generating knowledge about evaluation practice, it follows that a pragmatic perspective accepts that our worlds, and thus practices, continuously evolve through the day-to-day demands placed upon practitioners (Simpson, 2018). It views practice as continually evolving and, importantly, not simply an artefact of historical influences, but as a consequence of what we anticipate might happen (Simpson, 2018). As a result, this theoretical focus on the practitioner, and their practice, challenges the differences between what managers say they do and what they actually do (Mintzberg, 1973; Argyris and Schön, 1996).

Overall, the theoretical perspective on practice presented in the preceding paragraphs is consistent with the practitioner-oriented perspective I highlight in Chapter 2 where 'the practical' (Schwandt, 2003) reflects these ideas within pragmatism as a 'philosophy of practice'. These perspectives therefore reject a top-down view of theory or an evidence-based practice, rather that theory and practice should be mutually informing in a cyclical fashion (Leviton, 2015).

3.3 A pragmatic approach to conducting empirical research

As a result of the practitioner-oriented theoretical perspective, I was motivated to conduct research on evaluation practice using an approach that can be sympathetic to the goal of better aligning theory and practice. I do this through adopting a practitioner perspective on programme evaluation practice that accounts for the fact that evaluation practice is continually evolving as well as situated and social. I therefore adopt a pragmatic approach to research as a guiding framework for conducting the empirical component of this thesis. Below I discuss the reasons why I adopt a pragmatic approach.

A pragmatic approach to empirical inquiry sits outside of the traditional 'philosophy of knowledge' subfield of philosophy. Philosophy of knowledge deals with the issues of the nature of reality and truth, and offers a top-down approach to defining empirical inquiry, starting with ontological assumptions about the nature of reality (Morgan, 2007). This top-down approach then places constraints on the nature of knowledge and what can be 'known' from empirical research (Morgan, 2007). Moreover, the top-down approach of a philosophy of knowledge contributes to what Kuhn defines as the 'incommensurability of paradigms' (Kuhn cited in: Morgan, 2007) which calls into question the ability of researchers within different paradigms to communicate with one another. This is a consequence of the fact that researchers working in different paradigms are trying to access a different 'reality'.

In contrast to more traditional philosophy of knowledge approaches, a pragmatic approach shifts the central focus of methodological choices away from metaphysical questions of the nature of reality to questions of how and why researchers make choices about the way that they do research (Morgan, 2014). In terms of programme evaluation practice, this enables me to acknowledge, as I do in Chapter 2, that the various conceptualisations of evaluation practice are not necessarily incommensurable or mutually

exclusive, rather they simply reflect different sides of the same coin. This means that the different views on evaluation practice simply reflect one aspect of that practice rather than one perspective or another better reflecting the nature of evaluation practice reality.

In moving away from questions about the nature of reality, a pragmatic approach focuses instead on how the beliefs of researchers inform their actions through the process of 'inquiry' (Dewey, cited in: Morgan, 2014). Dewey defended this approach based on his belief in the idea of 'transactional realism' (Biesta, 2009). Transactional realism reflects the idea that we are always already in interaction (or transaction) with the world and reality, meaning that there is no gap between us and the world, or between us as researchers and 'reality': this stands in contrast to a more dualistic focus on the differences between mind and matter (Biesta, 2009). In transactional realism, it follows that we can only 'know' about the world as "a function of the ways in which we manipulate, interact, and intervene in the world" (Biesta, 2009, p.37): our knowing about a phenomenon is developed through the relationships between our actions and their consequences through the process of inquiry rather than our ability to access some external reality. The process of inquiry, according to Dewey's perspective, involves 5 iterative stages with continuous cycles between a researcher's beliefs and actions:

1. Recognising a situation as problematic
2. Considering the difference it makes to define the problem one way rather than another
3. Developing a possible line of action as a response to the problem
4. Evaluating potential actions in terms of their likely consequences
5. Taking actions that are felt to be likely to address the problematic situation

The focus of pragmatic inquiry is therefore related to questions of why researchers do the things that they do rather than simply describing how

research is conducted. In this respect, Dewey emphasises the role of human experience in research such that research does not occur in a vacuum (Morgan, 2014). Dewey argues therefore that all research is social and contextually influenced, and, as such, results in knowledge that is fallible and probabilistic (Morgan, 2014). Dewey refers to the knowledge outcomes of research as 'warranted assertions' that are a result of taking action and experiencing the outcome (Morgan, 2014). Warranted assertions contribute to our 'knowing' about the phenomenon, rather than generating 'knowledge' about the phenomenon, as would be the goal with a traditional 'philosophy of knowledge' perspective. This means that we can consider 'warranted assertions' as imperfect. This contrasts with a traditional philosophy of knowledge perspective that would strive to generate knowledge that reflects some reality, according to the philosophical perspective taken. In terms of this research, a pragmatic approach therefore means that the findings, and emerging discussion, reflect 'warranted assertions' about programme evaluation practice based on how I have intervened in the world, rather than reflecting the 'objective reality' of evaluation practice overall. This aligns with the practitioner-oriented perspective on evaluation practice that sees practice as continually evolving, as well as social and situated, and shaped by the lived experience of knowers. Therefore, any knowledge about practice is naturally contingent on those factors, and is thus imperfect.

3.3.1 Approach to inquiry

Taking the 5 stages of inquiry outlined above, I describe the process I followed in this research from recognising a situation as problematic to taking action that I felt likely to address the problematic situation.

Stage (1), Stage (2), and Stage (3): The literature review in Chapter 1 acknowledges that there are various ways to conceptualise evaluation practice that will affect how research on evaluation practice is conducted, what the focus of analysis is, and the outcomes of that research. That is not to say that

the differing conceptualisations are mutually exclusive, rather that different perspectives on evaluation practice are required to more fully understand an evaluation theory's use in practice from different perspectives and in different evaluation practice settings. These stages were also informed by my personal experiences of evaluation practice in small TSOs that did not quite align with what I was reading in the evaluation literature on programme theory. I therefore use Chapter 2 in this thesis to explore the different perspectives on evaluation practice in more detail.

Specifically, I discuss how the conceptualisations of the use of programme theory is predominantly focused on methodological application of programme theory by trained evaluators in systematic evaluation activities. Yet, this does not resonate with the image of evaluation practice in third sector organisations in both the literature and that which I had experienced in my interactions with such organisations. Therefore, I define 2 research questions to explore the use of programme theory in small TSOs that places the evaluation practitioner, rather than the method, as the focus of analysis.

Stage (2), Stage (3), and Stage (4): The current chapter focuses on defining a methodological framework to address the 2 research questions. Given that the research questions are respectively descriptive and prescriptive in nature, I discuss the value in adopting a multi-methodological framework that addresses each research question, using different methods that allow practitioners to both reflect on action and reflect in action.

Stage (5): the empirical component of this thesis is aimed at addressing the research questions in ways which I feel are appropriate to the generation of 'warranted assertions' about the use of programme theory in evaluation practice in small TSOs. By generating 'warranted assertions' I do not claim to generate knowledge about evaluation practice, in the traditional 'philosophy of

knowledge' sense, rather I contribute to our 'knowing' about evaluation practice by showing that this knowing is incomplete and context dependent. In adopting a pragmatic approach to generate 'warranted assertions', I am therefore acknowledging that research on evaluation practice is social and contextually dependent such that,

> *"Our experiences in the world are necessarily constrained by the nature of that world; on the other hand, our understanding of the world is inherently limited to our interpretations of our experiences. We are not free to believe anything we want about the world if we care about the consequences of acting on those beliefs"* (Morgan, 2014, p.1048).

However, It is worth noting that many utilise the pragmatic perspective due to its practical, rather than its theoretical and philosophical principles and values that I have discussed above (Morgan, 2014; Simpson, 2018). Morgan argues that pragmatism is often adopted because it seemingly advocates for a 'what works' approach, i.e. choosing the methods that work best for the research question. Whilst this is intuitively appealing to practically-minded researchers, it ignores the value of pragmatism as a 'theory of truth' that goes beyond simply choosing a methodological approach that suits the research question (Denzin, 2012, p.81). That is not to ignore the associations between a pragmatic perspective and problem solving per se, however it emphasises that the associations with methodological choices are not what defines the value of a pragmatic perspective in research (Morgan, 2014).

3.4 A multi-methodological framework

This thesis addresses two research questions with respect to the use of programme theory in programme evaluation practice in small TSOs. Given that these questions are respectively descriptive and prescriptive in nature I adopt a multi-methodological approach in order to address both research questions using appropriate methods. There are several reasons why I adopt a multi-methodological framework which relate to the value in affording practitioners the ability to reflect-on and reflect-in action, in addressing a research problem

from different angles, the ability to generate ‘warranted assertions’ through the process of the integration of findings from multiple methods. I have summarised the overall methodological approach to addressing the research questions in Figure 4. In the following sections, I describe the component parts of Figure 4 in more detail. First however I define and justify the multi-methodological framework adopted.

From a practitioner perspective, as outlined earlier in this chapter, Schön states that there are two ways which practitioners can ‘know’, or learn, about their practice, which allow research to access knowledge and learning about practice: that is through retrospectively reflecting-on-action (also referred to as ‘know-that’) and reflecting-in-action, where practitioners have access to tacit, or implicit, knowledge which is in some cases difficult to articulate (also referred to as ‘know-how’) (Schön, 1991; McAdam, Mason and McCrory, 2007). The methods adopted within the multi-methodological framework therefore place the lived experience of practitioners as the central focus of inquiry through facilitating the opportunity for practitioners to reflect-on-action and reflect-in-action. In doing so, this multi-methodological framework enables the empirical component of this thesis to address each research question respectively.

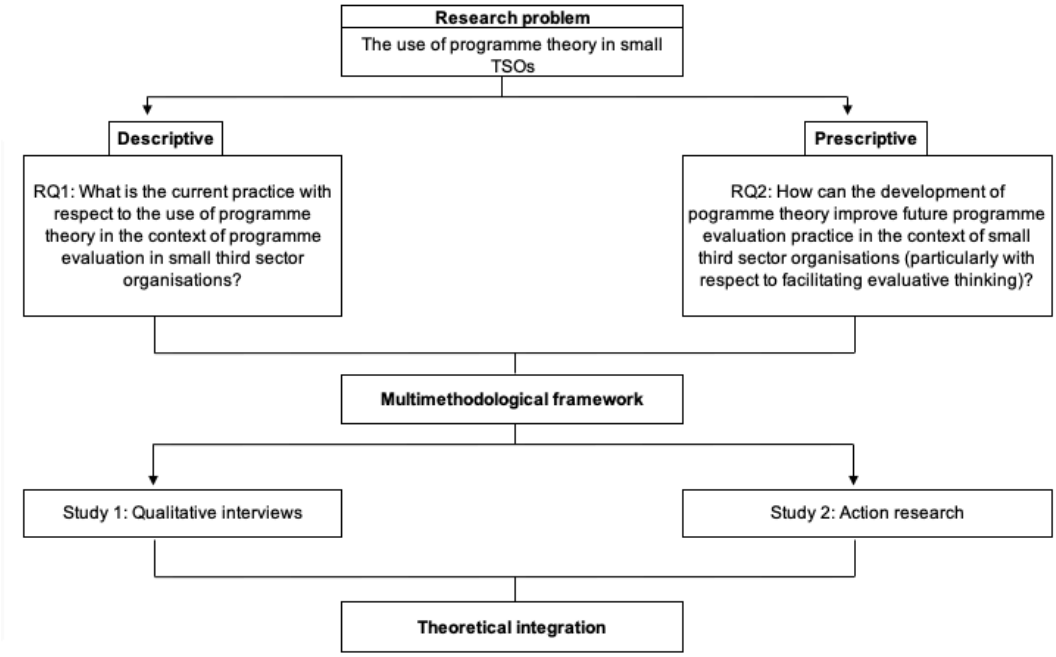


Figure 4 Methodological framework

A multi-method approach also allows research to address one research question or problem from different angles (Brewer and Hunter, 2006). However, I found some confusion in the methodological literature on the differences between mixed-method and multi-method research. For example, Hesse-Biber (2015) considers multi-method and mixed method research (MMMR) together in the opening chapter of a book on the topic. Nonetheless, for the purposes of this research I define mixed-methods research to refer to the use of a mixture of qualitative and quantitative methods in either a sequential or concurrent format to answer a research question. Multi-method research on the other hand, I define to refer to the use of multiple approaches to address a research problem from different angles with the aim of integrating the findings to form a more holistic picture of a phenomenon (Moran-Ellis et al., 2006). In this research, I adopt a multi-method approach in this sense to address the descriptive and prescriptive components of the research problem using different methods.

Generally, a multi-method approach also aims to address a major challenge facing social sciences: that is, any study employing one type of research method or philosophy of knowledge perspective leaves uncertainty surrounding other possible interpretations of data or hypotheses that might in some cases question the validity of the study's findings due to the weaknesses in any one method or limitations of any one philosophical perspective (Brewer and Hunter, 2006). The emphasis in social sciences on the use of multi-method approaches has emerged because "there is a strong tendency in all fields of social science for particular methods to be valued so highly … that they become ends in themselves, to be defended against rival methods and nourished by selecting only research problems for which they are well suited" (Brewer and Hunter, 2006, p.9) rather than addressing the research needs in question. In the case of the empirical component of this thesis, a multi-method approach enables me to utilise different methods to address two different research questions from a different perspective than that which dominates the literature on the use of programme theory in programme evaluation.

The overall research problem addressed in this thesis relates to the use of programme theory in programme evaluation in the context of small TSOs. The research questions addressed in this thesis relate to both understanding current practice as well as understanding how practice can be improved. The question on current practice is descriptive in nature as it seeks to describe how programme theory is used currently including the contextual influences on its use. On the other hand, the second research question addressed is prescriptive in nature as it seeks to address how the development of programme theory could improve programme evaluation practice in the future. I argue therefore that the differing nature of these research questions means that different methods can be used to address them effectively. In the following section, I describe the advantages to addressing both descriptive and prescriptive questions separately before detailing the specific methods I used to do so. The idea is that the descriptive and prescriptive studies will act as complements to one another in contributing to our 'knowing' about the use of

programme theory in programme evaluation practice in small TSOs through generating ‘warranted assertions’.

3.4.1 Descriptive study

The first research question addressed is a descriptive question of the current practice in small TSOs with respect to the use of programme theory. The aim of conducting a descriptive study is to give a sense of the context of programme evaluation practice in small TSOs with respect to the use of programme theory including current programme evaluation practices, the contextual influences on programme evaluation practice, and the extent to which programme theory is used within that practice.

However, a descriptive account can only address one aspect of this research problem. Argyris, for example, notes that often descriptive research unintentionally covers up gaps and inconsistencies in theory through its inability to change practice and contribute to understanding of those change processes and the influences, both positive and negative, that change can have (2003). Overall, this descriptive study aims to provide understanding, if only descriptively, of evaluation practice and its challenges as perceived by those practising evaluation which can also contextualise the learning from the prescriptive study. In this sense this study contributes to the evaluation field through a study which supports a dialogue between descriptive and prescriptive theory.

3.4.2 Prescriptive study

The second research question addressed is a prescriptive question of how the development and use of programme theory can help to improve future programme evaluation practice in small TSOs, particularly with respect to facilitating evaluative thinking. The aim of conducting a prescriptive study, as a complement to a descriptive study, is to generate what Argyris refers to as

'actionable knowledge' (Argyris, 2003): this means that this study aims to change something and observe that change. Overall, this prescriptive study aims to enhance and challenge our understanding of the *potential* role of developing programme theory in improving programme evaluation practice in small TSOs.

3.4.3 Integration through triangulation

An important task, in undertaking multi-methodological research, is in how the research will produce outcomes and learning resulting in an original contribution to knowledge (or 'knowing' in the case of a pragmatic approach). In multi-methodological research this is done through the process of integration which involves the triangulation of findings. In the case of this research, it entailed a process of integrating the respective findings of the descriptive and prescriptive studies to inform an effective dialogue between the two studies. Triangulation is "an epistemological claim concerning what more can be known about a phenomenon when the findings from data generated by two or more methods are brought together" (Moran-Ellis et al., 2006, p.47). The process of triangulation is central to integration in a multi-methodological study, both on a theoretical and practical level. In order to successfully integrate methods, Moran-Ellis and colleagues point to three different ways in which the process of integration can occur in a multi-method study:

1. Integration of methods: integration of methods is a term used only when methods have been integrated from earliest point of conception in a study and remain integrated throughout the process of the study. Integration is therefore achieved at both the theoretical and practical level from the outset of a study.
2. Separate methods, integrated analysis: this involves the common analysis of a diverse set of data without losing the distinct characteristics of the datasets. This can involve separate analysis within

each dataset and its paradigm, followed by the choice of an analytic question and following this question through the different data sets

3. Separate methods, separate analysis, theoretical integration: this involves the integration of the findings of each data set at the point of theoretical interpretation. This does not combine methods or analysis but focuses on bringing together findings in one interpretive explanatory framework. This type of integration can present challenges in terms of what to do with divergent findings but can at the same time value divergent findings within a pluralist framework.

In this study, I employ the third approach to integration i.e. separate methods, separate analysis, theoretical integration. This means that I conducted two separate empirical studies, one addressing the descriptive component and the second addressing the prescriptive component. Each study has, as described above, individual value and generates its own learning with respect to the descriptive and prescriptive questions. To triangulate these findings, I utilise the discussion chapter of this thesis to think more critically about the role and use of programme theory in programme evaluation practice in small TSOs in terms of the emergent learning from the two empirical studies. Whilst integration in this study occurs at the final stages of the research, I have underpinned both studies with the same principles and ideas, e.g. a practitioner perspective, however I considered it advantageous to use separate data collection and analysis procedures to address the descriptive and prescriptive components of the research problem separately.

3.4.4 Ethical considerations

Given the nature of what I describe in the following sections, ethical approval was applied for, and approved, via the University of Strathclyde Business School ethics committee. Ethics applications for the descriptive and prescriptive components of this thesis were approved separately. Moreover, in the case of the work with young people, I applied and was approved under the

PVG disclosure scheme. Other ethical considerations are discussed in the following sections.

3.5 Methods

I now describe the methods I used in two empirical studies which address the descriptive and prescriptive research questions respectively. First, I describe how I utilised qualitative semi-structured interviews to generate learning about current evaluation practice in small TSOs with particular attention to the use of programme theory. Second, I describe how I utilised an action research approach, and SODA (Strategic Options Development and Analysis) mapping, to generate learning about how the development of programme theory can improve future programme evaluation practice in small TSOs, particularly with respect to facilitating evaluative thinking. With respect to each method, I describe the method used, alternative methods considered, and how the data was collected and analysed. The final part of this section considers how I went about the process of 'theoretical integration'.

3.5.1 Descriptive study - qualitative interviewing

3.5.1.1 Justification of method – semi-structured interviews

In keeping with the empirical focus on practitioner experience and practitioner knowing, I adopt a qualitative approach to focus on the subjective perceptions and experiences (Miles and Huberman, 1994; Saunders, Lewis and Thornhill, 2016) of those practising evaluation in small TSOs. From the subjective perceptions and experiences of those practising evaluation, I aim to understand how programme evaluation is practised: this includes understanding the contextual influences on programme evaluation practice, and how and in what form programme theory is used. This involves exploring the practitioners' implicit theories of evaluation practice, particularly with respect to the use of programme theory. In order to do so I employ semi-structured interviews as the primary mode of data collection to elicit subjective perspectives and experiences.

Given that there were several topics I was interested in addressing through the interviews, the semi-structured interview format allowed me to address the primary topics of interest while allowing the interview format to be flexible. This meant that the order of and addition of questions, the terminology used, and the omission of questions where appropriate (Bryman, 2016) were all flexible. For example, given that programme evaluation practice in small TSOs is likely influenced by the implicit theories of those practising evaluation, the semi-structured format allowed me to explore the use of programme theory without having to use that specific terminology. As such, this allowed me to be sensitive to the ways in which the interviewees framed and understood issues, as well as being responsive to the direction which the interview took and the issues emerging from responses to each question (Bryman, 2016). The structure, in terms of topics addressed, paired with a degree of flexibility, facilitated the flow of conversation in each interview setting (Saunders, Lewis and Thornhill, 2016). Moreover, in terms of generating learning about current practice, the use of semi-structured interviews afforded participants the ability to reflect-on-action which involves retrospectively reflecting after an event(s) has occurred and involves reviewing, analysing, and evaluating a situation (Schön, 1991). Interviewees often commented that this was the first time they had the chance to sit down and talk about evaluation.

On a practical level, there are several advantages to using the semi-structured interview method and qualitative methodologies more generally in the context of third sector organisations. Soteri-Procter (2010) identifies several advantages of qualitative methodologies in such settings including: the flexibility of qualitative tools and the ability to gain information on processes and on staff perspective (2010, p.414) and the ability of qualitative approaches to facilitate building rapport and communication with respondents (ibid. 2010). Building rapport and communication with respondents was important in this study where participants were generous with their valuable time and also to

ensure that participants felt comfortable in openly speaking to someone they do not know very well about their experiences with programme evaluation; this often presents a challenge and can be a source of stress for 'accidental evaluators'.

Alternative methods

Semi-structured interviews are not the only method for conducting descriptive research. A survey was also a viable option had the literature provided more conceptual understanding of the role of programme theory in the context of evaluation in small TSOs. Had it been clearer for example from the literature, how or in what forms programme theory was used by accidental evaluators, a survey would have enabled a wider population of evaluation practitioners in small TSOs to be contacted to understand the use of programme theory more widely. However, a survey format may have been difficult given this lack of conceptual understanding and the fact that the participants are accidental evaluators often implementing their own implicit theories of evaluation practice and as such may be unfamiliar with specific terminologies necessary to construct survey items e.g. 'programme theory'. Moreover, even with the use of open-ended items in a survey, it may have been difficult to elicit the same information as can be elicited from the depth of discussion facilitated by a qualitative interview format, especially with respect to understanding the evaluation practice context.

3.5.1.2 Selection of participants

The population of interest in this study are those practising evaluation (as accidental evaluators) in small TSOs delivering programmes and services to address social issues. In the UK for example, small TSOs make up, in number, the majority of third sector organisations[15], and therefore, represent a large

[15] https://www.ncvo.org.uk/about-us/media-centre/briefings/1721-fast-facts-about-the-charity-sector

population of organisations with a presence in local communities across the country. I was also interested in gathering the perspectives from those within organisations which might influence programme evaluation activity, or the context of programme evaluation practice within small TSOs. I identify these organisations as funding bodies or trusts, i.e. those who fund third sector activity, and evaluation support organisations, i.e. those organisations who play a role in training and capacity building with respect to programme evaluation activity within small TSOs.

Miles and Huberman note that qualitative sampling is not done to be representative of the larger population, rather to gain an in-depth understanding for some conceptual or theoretical reason (Miles and Huberman, 1994). For practical reasons and in order to obtain a bounded sample of organisations, I initially requested a list of small TSOs delivering services to children and young people through Glasgow Council for the Voluntary Sector (GCVS) *Infobase*[16] service. GCVS is a membership organisation open to all third sector organisations that acts as both a development agency and advocate for voluntary and community organisations in Glasgow[17].

This request generated a list of over 385 organisations. From this list, I focussed on those organisations addressing social issues, excluding organisations on the list with annual incomes over £1.5million pounds[18], and those which did not have a focus on children and young people. Whilst I was

[16] http://www.infobaseglasgow.org – Infobase is a database of over 1500 community and voluntary organisations

[17] http://www.gcvs.org.uk

[18] The Small Charities Coalition (https://www.smallcharities.org.uk/who-we-are/) defines small charities as those with an annual income of under £1million. I chose £1.5 million as a defining amount to widen the scope of participant recruitment slightly. However, I recognise that this is only an indicative indicator and thus not wholly reflective of the 'size' of an organisation.

primarily interested in the concept of ‘accidental evaluators’, it was difficult to use this as my defining term to identify research participants. I therefore used the income level defined above as an initial indication of the size of the organisation, and thus the likelihood of there being an accidental evaluator. This follows a purposive sampling approach, i.e. “selecting respondents based on their ability to provide the needed information” (Padgett, 2008, p.53). This produced a list of 126 organisations, all of whom were contacted initially via e-mail, and, where contact numbers were available, organisations were subsequently contacted via phone call due to bounced e-mails, to gauge their interest in participating in the study. I confirmed with those willing to participate, that those conducting programme evaluation were ‘accidental evaluators’. This approach yielded a low response rate of roughly 4 organisations indicating willingness to participate. This is unsurprising given the small size of organisations, many lacking resources for core programme administration, and others relying primarily on volunteers. It cannot be expected in such cases that potential participants would be able to afford an hour out of their day to participate in research, additional to their own daily demands.

Following this, and based on the initial four interviews, a snowball sampling approach was used to identify other organisations to interview. Snowball sampling is a type of purposive sampling used to identify participants that are difficult to identify or contact (Padgett, 2008). I conducted snowball sampling by asking whether those I initially interviewed knew of any other organisations with staff who might be interested in participating. This approach yielded a further 8 organisations with participants willing to participate. Also willing to participate were 3 funding bodies and 2 evaluation support organisations; participants from these organisations ensured that the interviews captured insights from those organisations which influence and support evaluation practices in small TSOs.

Of course, participation bias is reflected in the nature of participants who choose to participate in the study (Saunders, Lewis and Thornhill, 2016). It may be that some level of participant bias exists in the study in that those who have chosen to participate could be those who feel strongly about programme evaluation practice in one way or another. Whilst I do not necessarily consider it to be a bad thing that interviewees were potentially enthusiastic to talk about their experiences with evaluation (positively or negatively), I was mindful of this in writing up the findings of this study, i.e. that findings are only reflective of those who chose to participate rather than reflecting the sector's perspective as a whole.

As I was interested in organisations that are small in size (based on income), this often meant that there was only 1 full-time member of staff available for interview. However, in cases where 2 members of staff were able to contribute, multiple interviews were carried out within respective organisations. Having multiple perspectives enabled a richer understanding of the experience of programme evaluation within that organisation and avoids what Carman refers to as the 'ecological fallacy' of having only 1 perspective to represent multiple perspectives within an organisation (2011): this was the case for 5 of the TSOs. In total, 17 interviews were carried out within third sector organisations with an additional 6 participants from third sector funding and evaluation support organisations. Table 4 provides a summary of the types of organisations, role of interviewees and number of interviews conducted per organisation.

Organisation code	Focus of organisation	Role of interviewees	# interviews conducted
OrgA	Youth/community development	(1)Project manager (2)Youth/health & wellbeing officer	2
OrgB	Youth project	(1)Project manager	1
OrgC	Youth with complex needs	(1)Project manager (2)Development officer	2
OrgD	Youth project	(1)Project co-ordinator (2)Detached youth worker	2

OrgE	Youth project/youth employment	(1)Project manager	1
OrgF	Community development	(1)Programme co-ordinator	1
OrgG	Disadvantaged young women	(1)Board member	1
OrgH	Community development	(1)Community development officer (2)Community development worker	2
OrgJ	Disadvantaged young women	(1)CEO	1
OrgK	Youth employment	(1)Compliance manager	1
OrgL	Youth project	(1)Executive director (2)Director	2
OrgM	Youth employment	(1)Community representative (activist)	1
SupOrg1	National evaluation support organisation	(1)Training officer	1
SupOrg1	Consultancy evaluation support organisation	(1)Director	1
Fund1	Cross-UK funder	(1)Evaluation officer	1
Fund2	Local funder	(1)Evaluation officer (2)Evidence & learning officer	2
Fund3	Cross-UK/international funder	(1)National head (Scotland)	1
		Total:	**23**

Table 4 Sample organisation types, participant roles, and number of interviews conducted

3.5.1.3 Data collection – preparing for and conducting the interviews

Development of interview question guide

Overall, this study is focused on understanding the use of programme theory in current programme evaluation practice in small TSOs from the perspective of those practising evaluation. My reading of the literature on doing evaluation revealed several important aspects of programme evaluation practice which were incorporated into the question guide. These are the capacity to do evaluation, programme evaluation activities, and the purposes and use of evaluation. The question guide also included questions about the contextual factors influencing programme evaluation practice. Primarily, I was interested in the use of programme theory so this was also incorporated in the question guide. However, I acknowledged that familiarity with this concept would be varied and this, I recognised that I might have to draw out the use of programme theory myself, rather than ask participants directly about it. The

questions topics, relating to the research question, included in the question guide are illustrated in Figure 5.

Reliability concerns in qualitative research relate to the lack of standardisation in approach, and how this might affect the ability of another researcher to generate similar findings (Saunders, Lewis and Thornhill, 2016). I argue however that the goal of this study was to capture the dynamic nature of programme evaluation practice as well as contextual influences on this process, which is seen to be a strength of qualitative research. More standardised research designs, such as survey research, may well undermine this goal of capturing rich information. The semi-structured interview format enabled me to use a question guide in each interview (Appendix 2) to ensure that each interview addressed the same questions and topics with respect to the current programme evaluation practice and use of programme theory.

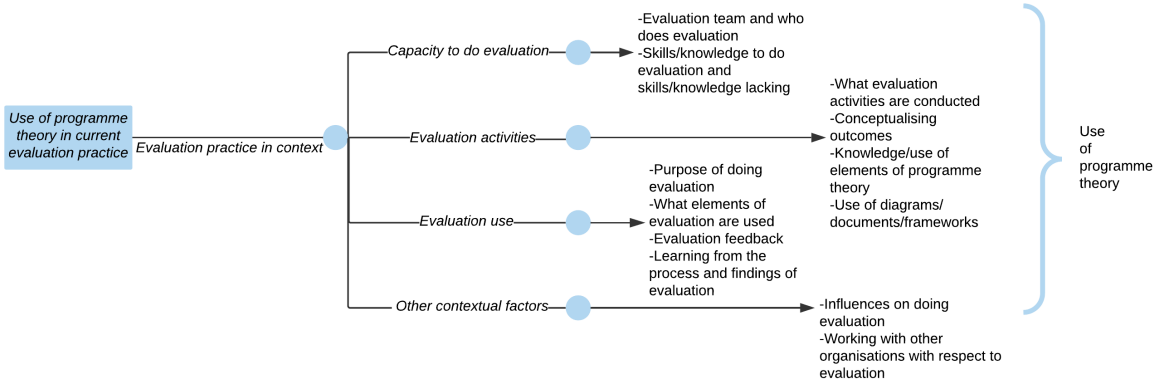


Figure 5 Links between RQ and interview guide question topics

Moreover, the way an interviewer conducts the interview, including tone, questioning approach, comments, and non-verbal behaviour, can introduce bias, influencing the way in which the interviewee responds (Saunders, Lewis and Thornhill, 2016). The use of a question guide enabled me to have a specific format to stick to so that the use of irrelevant questions or comments was minimal. Whilst I do not think a question guide can always control for use

of tone and non-verbal behaviour, I do think it provided sufficient structure to help deal with this form of bias. Overall, I tried to use a tone in the interviews that was friendly and conversational, whilst adhering to the question guide. I did not want to come across as interrogative and wanted to ensure that participants felt like I was appreciative of the valuable time that they were giving me.

To ensure that the interview format was both accessible for the interviewee, as well as a reliable tool for data collection, a draft question guide was piloted with 2 participants in an organisation not to be included within the main study. The first draft of interview questions centred on understanding of programme theory and its use within programme evaluation activity. However, what became clear quickly was that the term programme theory was unfamiliar and, as a result, some of the questions became difficult to answer for the participants. Whilst explanation cleared up misunderstandings, I felt it was inappropriate to continue with such a specific format, rather I needed to be more sensitive to the fact that participants were 'accidental evaluators' and that their implicit theories of evaluation practice might not align with the terminology used in the pilot interviews.

Having completed the pilot interviews, the question guide (Appendix 2) was updated to focus more broadly on programme evaluation activities at a more general level, out of which I could conclude the extent to which the principles of programme theory were currently being used. The second draft of the interview protocol intended to facilitate a broader understanding of the context of programme evaluation practice and thus provided a better means of addressing the research question. The second draft was structured to ensure interviewees felt comfortable and that the question format was accessible. To ensure this, the question guide followed the format laid out in Creswell and Creswell (2018). The structure of the question guide included the following:

1) Basic information about the interview, what it involved and the opportunity to ask any questions before beginning the interview offered participants the opportunity to clarify any uncertainties.
2) An introduction to the researcher allowed the interviewees to both feel more comfortable with the researcher and to give the interviewees some context about where the motivations for the study was coming from
3) Opening questions in 'ice-breaker' style helped to get the participants to talk about themselves, their organisations, and programme evaluation in a way that is accessible
4) Content questions (5-10) with probes which included the main topics/questions to be address in the interview e.g. with respect to programme evaluation practice (capacity, activities, and use) and the contextual influences on evaluation practice, as well as understanding the use of programme theory
5) Closing instructions which included asking if the interviewees had any questions for me, thanking the interviewees for their time, informing the interviewees about my next steps, and informing them about how to contact me should they wish to do so

Interviews (timing & location, and pre-, during, and post-interview contact)

Interview location and timing

Interview location was primarily chosen for its convenience to the participant. In the majority of cases, this meant that interviews took place at the offices of the participants' organisation and at a time which was mutually agreed on between myself and the interviewees. In one instance, the interview was carried out via telephone call due to the interviewee working from home.

Pre-interview

To minimise bias in the conduct of interviews, all participants were informed about what would be expected of them in the interview. I ensured that all participants received a participant information sheet (Appendix 5) and consent form prior to the interview taking place. This informed the participants about the goals of the study and the kinds of information that would be required of them during the interview. Moreover, via e-mail or phone call, participants were given the opportunity to clarify any questions regarding the research generally or about the interview format. I also gave participants the opportunity to ask questions directly before and after the interview had taken place. Before the interview, I ensured that each participant was able to review and sign a consent form which indicated my data collection, analysis, storage, and management plans and allowed them to indicate their willingness to be audio recorded. In all but one case were the interviews were audio recorded. In the case where the interview was not audio recorded, detailed notes were made during the interview.

During the interview

An important factor to consider with respect to carrying out the interviews is when the interviewee chooses not to reveal their experiences in some ways so that the researcher can only obtain a partial picture (Saunders, Lewis and Thornhill, 2016): this, paired with a poorly defined question guide, can affect the validity and credibility of the study's findings where validity and credibility refer to the researcher's ability to gain access to a participant's knowledge and experience and be able to infer the meaning intended by each participant (Saunders, Lewis and Thornhill, 2016). To promote the validity within this study, the pilot interview served as a basis for me to understand appropriate question formats within this setting. By using a combination of open questions, probing questions and specific questions, I was able, to the best of my ability, to gain access to the participants' subjective perspectives and experience. I feel as though the use of open questions, to begin with, allowed the

participants to set the scene in terms that were familiar to them. For example, the interviews began with the questions:

> *"Tell me about your organisation"* and*,*
> *"Tell me about your approach to programme evaluation"*

These questions enabled the participants to speak about their organisation and approach to programme evaluation using whatever terms that they wanted and provided me with a sense of the language used and how I should proceed with the interview.

Another cause for concern, with respect to validity and credibility, is the extent to which the participants *can* reflect-on-action in a way which reflects what actually happens in practice. For example, there is often a disconnect between what people say and what they do i.e. a difference between their 'espoused theory' and their 'theory-in-use' (Argyris and Schön, 1996). This is an issue which is difficult for a descriptive study such as this to address. By using open question formats, purposefully designed to be answered in a way which was accessible and meaningful for the participants, allowed them to speak about programme evaluation in a way which was comfortable to them. Probing questions were then used to ask for more detail from specific responses. An example of this was when participants talked about activities that, from my perspective, were inherently evaluative, but the participant did not consider or perceive them to be 'programme evaluation'. In such cases I was able to use questions such as, 'can you tell me a bit more about that' to ask for more detail and to understand better the implicit theories and perceptions of programme evaluation.

Post-interview

As soon as possible after the interview, I consolidated any notes I had taken during the interview and reviewed any evaluation documents offered to me by the interviewee with respect to evaluation. Such documents included examples of reporting templates, evaluation questionnaires, and examples of evaluation documents/reports. Whilst these were not used in the formal analysis, because they were not available or offered by every participant or they could not be shared, they did provide some context and confirmed much of what was discussed in the respective interviews.

In addition to this, audio recordings were transcribed verbatim. I could argue that note taking from audio recordings would have been sufficient; however, I found that the rich data and context available in the audios meant that full verbatim transcription in the first instance, gave me a first chance to get to know the data. Second, in transcribing verbatim and using this to analyse the data, I did not risk losing the context of what was said by participants if I had simply taken notes from the audio recording. This was particularly important given the focus on eliciting 'implicit theories' of programme evaluation practice. I therefore had to be mindful of the context in which things were said in case participants were referring to programme evaluation activities that they did not label as such for example.

3.5.1.4 Data analysis - coding and thematic analysis

To analyse the data, I employed thematic analysis. This involved several iterative phases of reviewing data (transcripts), coding the transcripts, and developing themes based on the codes. Creswell and Creswell note several steps to be undertaken whilst conducting thematic analysis (2018) which, in the context of this research, I list and discuss below:

1) Organise and prepare the data: this involved the post-interview note consolidation, transcription process, and filing both notes,

transcriptions, and any other materials given to me by interviewees in a file. This meant that all materials were easy to review together. I found that having the material in print facilitated the analysis at an early stage by being able to work with pen-and-paper in a 'messy' way rather than in computer software. This was a personal choice. I explored the use of NVivo qualitative data analysis software but I felt less comfortable working with rich data from a computer screen. I liked being able to move between my own notes and the transcripts. The process of working with pen and paper facilitated my own thinking processes which I struggle with when using a computer screen.

2) Read and review all of the data: once transcriptions were individually complete, I usually spent some time reading through them to get a sense of the information from each interview and reflect on the data. Once all interviews were transcribed and filed, I reviewed all transcriptions together and took some notes on some codes I thought might be useful to begin with.
3) Start coding the data: I used an iterative approach to attaching codes to data. In reviewing each transcript, I noted potential codes in the margins. Codes referred to broad topics in the interview including: evaluation practices (capacity, activities, use), factors influencing evaluation practice (i.e. contextual factors), and the use of programme theory. I used a 'template approach' whereby codes can be 'directly imported from the interview questions with varying degrees of openness to new information and new codes' (Patton, 1990; Padgett, 2008).

Framework matrices were used as a tool to gather the data from all interviews using the 'template approach' by presenting data in a table of rows and columns where rows represent cases and columns represent themes and sub-themes using the codes. Displaying data in this way allowed me to think about research questions in a more focussed way, including identifying which data is relevant to answer the research question; it allowed me to consider the data more holistically,

keeping all potentially relevant material whilst structuring the data in a coherent and systematic way (Miles and Huberman, 1994).

Microsoft Excel was used to tabulate the data. As a result, emerging themes could also be identified under each broad heading. Interview data was coded 'in vivo' by 'transferring chunks of text into conceptual bins' (Padgett, 2008, p.151), the bins referring to the codes and emerging themes. Data entry into the matrix followed an 'in vivo' approach to maintain clarity and detail in the earlier stages. This is what Miles and Huberman refer to as using 'thick' entries (Miles and Huberman, 1994) which helped me to retain the context of what was said by participants.

I used blank Microsoft Word documents to collate all the 'thick' entries relating to the broad headings in the Excel document (capacity to do evaluation, evaluation activities, evaluation use, contextual influences, and use of programme theory). This facilitated a process of coding specific to each broad theme which would then inform the development of more discrete sub-themes. A coding hierarchy which summarises the resulting themes and sub-themes in shown in Appendix 3.

4) Generate a description and themes: using the Microsoft Word documents alongside the emerging codes, I generated descriptions of each sub-theme, where appropriate. This facilitated a re-organisation of the data into a format which assisted the final stages of analysis whereby I looked more specifically at each theme and sub-theme in isolation and considered any cross-linkage between themes.
5) Write-up: this involved drafting narratives about the themes and sub-themes, utilising quotations as evidence. This was an iterative process to arrive at the findings of this study.

3.5.2 Prescriptive study - action research

The second research question addressed in this research is related to how programme evaluation practice can be improved through the use of programme theory. Argyris advocates that actionable knowledge, such as understanding how to improve evaluation practice, can be generated through changing and challenging the status quo (Argyris, 2003) and observing what happens as a result. In order to address the question of how programme evaluation practice can be improved through the use of programme theory, I adopt an action research approach to develop programme theory within two organisations to understand the added value of doing so in the context of programme evaluation practice in small TSOs.

3.5.2.1 Justification of method – action research

I employ an action research approach to address the prescriptive question of how programme evaluation practice in third sector organisations can be improved though the use of programme theory. To do so, I employ the definition of action research given by Eden and Huxham (2002, p.254) that action research is an approach to *"researching organisations…which results from the involvement by a researcher with members of an organisation over a matter that is of genuine concern to them and in which there is an intent by the organisation members to take action based on the intervention"*. According to this definition, action research is defined as a collaborative approach to research that provides participants with the means to take action in order to change (Stringer, 2007). Action research does so through allowing participants the opportunity to reflect and act by carrying out an intervention in the organisation (Bradbury, 2015).

Consistent with a practitioner-oriented theoretical perspective, action research therefore places practice as the focus of inquiry in order to draw on and contribute to both practice and theory (Bradbury, 2015). In this way, action research combines the notions of theory and practice in contributing to our 'knowing' (Robertson, 2000; Bradbury, 2015). Robertson highlights the

importance of combining theory and practice through the concept of reciprocity in action research, the need to actively draw learning from but also contribute to action in the participants' lives, organisations, or communities (Robertson, 2000).

3.5.2.2 Scope of the action research study

There are however many different ways of implementing an action research approach (Herr and Anderson, 2015). Action research approaches range along a continuum from fully participatory approaches to empowerment and change (e.g. Participatory Action Research) through to action research approaches more closely aligned with applied research. The point I wish to make when defining my approach is how I strike a balance between generating learning i.e. conducting research, and generating action and change (Huxham and Vangen, 2003). Balances and trade-offs need to be made depending on the ultimate goals of the research and in particular with respect to the extent to which you can achieve both rigour and relevance in action research (Eden and Huxham, 1996; Huxham and Vangen, 2003; Morton, 1999).

As the goal of a PhD thesis is to contribute to the academic field, I justify the use of action research through defining how I use it to contribute to our 'knowing' about the research problem. Specifically, there are questions around the generality of action research findings i.e. the implications beyond the actions taken in the specific context of the action research (Eden and Huxham, 1996, 2002; Morton, 1999). First, this study addresses some of these concerns through the application of action research in two different organisations, rather than in a single organisation. Whilst there were specific actions taken in each organisation, the use of two empirical examples allowed me to think about the wider implications of and learning about those more specific actions.

Second, the primary goal of action research is not to test existing theory, rather to focus on the emerging theory and learning, and the development of existing theory (Eden and Huxham, 2002). Working with both organisations over time allowed me to engage in continual reflection on the role of the development of programme theory in both organisations through each cycle of involvement with the organisation (Eden and Huxham, 2002). This meant reflecting theoretically on the research question during problem definition, conducting meetings with the organisation, conducting the facilitated discussion/workshop within each organisation, reflecting on meetings and the facilitated discussion/workshop, and as a result adapting my learning and approach as I saw fit.

The core component of Eden and Huxham's definition of action research is the action research intervention which serves the purpose of contributing to change within the action research setting. The intervention in the case of this research is the development of programme theory. However, one of the challenges of conducting action research rigorously is the idea that such an intervention is a 'one-off' and as such, not repeatable. This can impinge the validity, particularly the external validity, of action research as an approach to generating learning (Eden and Huxham, 2002). Whilst action research does not seek to achieve rigour in the same way as traditional scientific method, I did try to ensure rigour by implementing an intervention underpinned by the same methodology but applied differently in the case of each organisation. The application of the same methodology in each case supported the ability to draw out common learning point across both applications through comparing similarities and differences whilst meeting the individual needs of each organisation.

The concern about external validity is challenged by advocates of action research who believe that theory about how to improve practice can be improved through generating 'actionable knowledge' (Argyris, 2003).

Actionable knowledge is generated through changing and challenging the status quo (Argyris, 2003) and observing what happens as a result. As Eden and Huxham state, *"reliable data, and hence theories, about both past and future aspects that influence the way in which people change a situation are much more likely to emerge from a research process that is geared to action than more traditional approaches"* (2002, p.264). Action research, and the approach I have taken, does this through the ability to observe what people actually say and do in situations and contexts that actually matter and have value to them, i.e. to observe the situated doings and sayings of programme evaluation practice (Simpson, 2018). This stands in contrast to what they might say they would do hypothetically or in a controlled experiment. In the words of Argyris and Schön, action research targets theories-in-use (what people actually do) rather than espoused theories (what people say they do or would do) (Argyris and Schön, 1996, 1992).

3.5.2.3 Alternative methods

Due to some of the criticisms of action research as a valid and rigorous approach to generating learning, there are alternative methods which could be considered to address a prescriptive research question. An example would be a controlled experiment where I implement a standardised intervention and observed what happened to the individuals I am observing. However, one reason for not implementing a controlled experiment in this setting is because there is no standardised intervention for developing programme theory in settings such as TSOs for the purposes which this research serves. There are certainly sets of guidelines for developing and using programme theory (see Appendix 1 for more details of such approaches) however, as the goal within these organisations was simply to develop programme theory, I choose an approach that would enable this in each organisation given the specific programme evaluation challenges they were facing. In the literature on programme theory, it is clear that the purpose of using the approach should define the approach taken and as such, I chose to use an approach which could be adapted rather than a standardised intervention.

3.5.2.4 Accessing the field

Contacts with each of the organisations participating in this action research study were developed in the first instance through university contacts. The organisations were primarily interested in developing links with the university to support their research and knowledge needs. The proposal for this study emerged from my sustained involvement in preliminary meetings with each organisation. These meetings were initially introductory for organisations, introducing them to the university and relevant people within the university. However, from this point, I was able to pursue more targeted contacts based on discussions about the challenges facing and research interests of each organisation. These challenges and research interests mostly centred on conducting evaluation. The organisations were concerned in some way with improving their internal capacity to do programme evaluation, rather than simply seeking research conducted by the university. It was however important at this point that I did not impose my research interests on the organisations, rather that it was their specific challenges and research interests that drove how we proceeded. This meant that the action research was addressing an actual need of the organisation, rather than me imposing a study on them.

3.5.2.5 Organisational characteristics

In Table 5, I have outlined some of the key organisational characteristics of the two organisation within which the action research study was conducted.

	Organisation 1	**Organisation 2**
Annual income (2016/17[19])	<£400,000	<£2.5 million overall (community arm of work £37,000)
Number of FT/PT staff	~15 programme staff (& ~15 school staff co-funded)	2 full-time staff
Type of organisation	Scottish Charitable Incorporated Organisation	Statutory corporation (with community arm)

[19] Estimate relating to the financial year prior to coming into contact/working with each organisation

Evaluation	Internal evaluation with accidental evaluation (some prior evaluation experience) Seeking external evaluation	Internal evaluation with accidental evaluators

Table 5 Action research - organisational characteristics

3.5.2.6 The action research intervention: the SODA methodology

As I have previously mentioned, I implemented an action research intervention in two organisations. In the following sections I seek to describe the methodology I used to develop programme theory and justify its relevance. I briefly describe how the methodology was applied in each organisation. However, I wish to use chapter 5, which presents the findings of this study, to detail how the specific approaches came about and were implemented as I consider the whole action research process (from problem definition, to identifying solutions, implementing solutions, and evaluating solutions) as an integral part of developing the emergent findings in action research (Bradbury, 2015). However, in what follows in this chapter, I describe and define the methodology which forms the basis of the action research intervention, briefly describe how it was implemented in each organisation, and finally describe the processes of data collection and analysis during and after the intervention in each organisation.

The intervention specific method adopted, but applied in different ways within two organisational settings, is the SODA (Strategic Options Development and Analysis) methodology. In each organisation I applied the SODA methodology as a tool to facilitate the development of programme theory according to the specific evaluation needs of that organisation. In what follows, I detail the aims of the SODA methodology and how this methodology aligns with the concept of developing programme theory.

Aims of SODA methodology

The SODA (Strategic Options Development and Analysis) methodology was developed within the field of strategy development (Ackermann and Eden, 2010) and aims to elicit and build directed graphs, or 'maps', of a problem or issue area within, or facing an organisation as a means to develop solutions and action for the given problem area. The SODA methodology is part of a family of approaches in the field of Operational Research referred to as Problem Structuring Methods (PSMs) which focus on being able to structure and frame a problem before thinking about how to 'solve' the problem (Rosenhead, 1996). Working with more ill-defined problems, Rosenhead (1996, p.119) characterises PSMs as: representing problem complexity graphically (rather than in tables for example); and as inclusive of multiple perspectives meaning there is greater exploration of the explanations of the problem than being constrained by only one or two potential explanations.

The idea is that the inclusive and interactive nature of PSMs can offer insights into complex problems which traditional methods of analysis are not amenable to by exploring in-depth potential problem explanations and solutions (Rosenhead, 1996). Such techniques are advantageous because: 'decision-makers are more likely to use a method and find it helpful if it accommodates multiple alternative perspectives, can facilitate negotiating a joint agenda, functions through interaction and iteration, and generates ownership of the problem formulation and its action implications through transparency of representation' (Rosenhead, 1996, p.119).

This methodology aligns with the process of developing programme theory in a number of ways. First, developing programme theory can often involve working with ill-defined problems, characterised by complexity, in that programmes can often involve multiple and interacting components (Stame, 2004; Rogers, 2008). Through the use of SODA maps, users are able to capture, work with and analyse potentially complex issues to develop strategic

actions to address such issues. Second, in many cases, there may be multiple and sometimes competing perspectives on the programme theory (Friedman, 2001). The participatory focus of the SODA methodology aims to incorporate multiple perspectives on a problem or strategic issue as well as acting as a tool for facilitated discussion and negotiation amongst stakeholders. Last, the visual nature of the maps aligns with other approaches to depicting programme theory where for example, a box and arrow diagram may be used. However, a criticism of many depictions of programme theory is that they are over simplified versions of a programme (Davies, 2018). An advantage of the use of SODA mapping in this sense is that complexity can be incorporated into the visual structure of the map with the option to simplify the visual through analyses of specific parts of the map.

I now discuss these ideas in more detail in emphasising their relevance to developing programme theory. In particular, I comment on the relevance of: map structure, the contents of maps, the analysis of map structure, and mapping processes in the SODA methodology.

Map Structure

The structure of SODA maps aligns with how to think about programme theory. SODA maps are means-end diagrams drawn as short pieces of text (10-12 words) linked by unidirectional arrows, representing believed causality (Eden, 1988; Eden, Ackermann and Cropper, 1992). A map can either be created on an individual basis, called a cognitive map, or a map reflecting the views of a group is referred to as a cause map as it is no longer an individual's thinking that is represented in the map (I discuss these types of maps in a later section on mapping processes). Each short piece of text represents a 'construct' i.e. one key idea or assertion about a problem or issue (Eden, 1988). Using verbs in each construct helps to ensure that the map maintains an action orientation through its means-end structure.

In most cases, the top of a map will represent the hierarchy of goals an organisation wishes to achieve with respect to the given problem or issue, with the strategies to achieve those goals represented below followed by more detailed options/constraints impacting those strategies (Eden, 2004). Those goals at the very top of the map with no arrows pointing out are called 'heads' and those more detailed options at the bottom of the map with no arrows pointing in are called 'tails'. An example of the generic structure of a cognitive map, adapted from Eden (2004) and Eden and Ackermann (2001), is shown in Figure 6; it illustrates the functionality of the arrow as well as the hierarchical means-end 'tear drop' structure of a cognitive map within the SODA methodology with both positive and negative goals at the top of the map (also called heads), with the most detailed options or strategies for achieving those goals positioned at the bottom of the map (also called tails). A highly central statement is likely to appear in the center of the map and is a statement which is likely to be highly influential on the map structure and meaning. This is often indicated by having the most arrows in and out of the central statement, meaning that the statement is influenced by the most statements as well as influencing the most statements. I argue that this means-end structure aligns well with the concept of programme theory where there is a desired change or programme outcome and potentially various means of achieving or influencing that outcome. Such means could relate to specific programme activities or the mechanisms generated through engagement in programme activities.

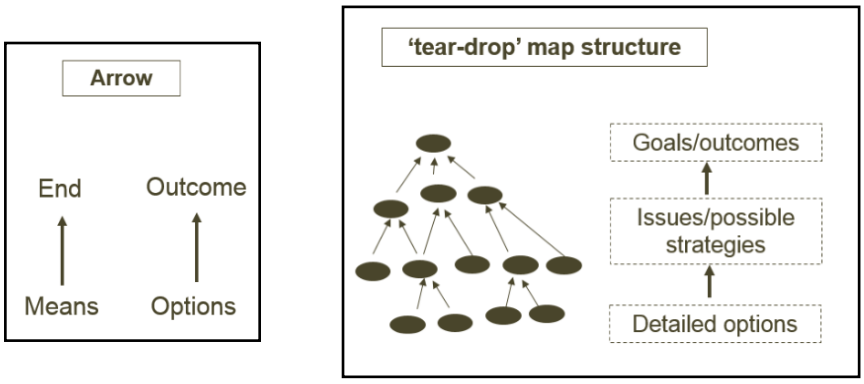


Figure 6 Cognitive mapping functionality - adapted from Eden (2004) and Eden and Ackermann (2001)

Map content

The content of SODA maps is also relevant to programme theory. Cognitive maps are developed by mapping a person's thinking about a problem or issue (Eden, 2004). Individual cognitive maps can, with care, be merged to form a group map (cause map). In many cases, cognitive mapping involves the use of open or semi-structured interviews (Eden, Ackermann and Cropper, 1992) to elicit the individual's thinking. Including content in the map in this way is relevant in the case of developing programme theory for the reasons outlined below.

Theoretically, cognitive mapping is based on Kelly's Theory of Personal Constructs (Eden, 1988) which argues that we continually try to make sense of the world to anticipate and reach out for the future. Cognitive mapping is based on three assertions from this theory: that one makes sense of one's world through contrast and similarity; that one seeks to explain ones world; and that one understands the significance of one's world by organising thinking hierarchically in that some concepts will be considered superordinate to others (Eden, 1988).

Maps are based on qualitative data representing subjective perspectives. The use of qualitative data in this way reflects the fact that those working in organisations think and work most of the time with language and ideas rather than variables, numbers, or mathematical models (Eden, 1988). With that, the SODA methodology aims to elicit subjective perspectives on a problem or issue area and explore those perspectives using map analysis (Eden, 2004). Through the incorporation of multiple subjective perspectives, the use of SODA mapping can broaden an issue area whilst managing complexity through the analysis of the structure of the maps (Eden, 1988). As a result, the use of the SODA methodology can help to elicit subjective perspectives on

the successful or unsuccessful functioning of a programme, often teasing out multiple perspectives or reasons.

In many cases, the understanding of how and why programmes are working to achieve desired outcomes may differ across programme stakeholders, whether that be programme management, those delivering the programme, or programme beneficiaries. Empirical studies have highlighted the importance of eliciting these differing perspectives to surface and challenge the varying assumptions about the programme before thinking about how to go about evaluating that programme (Friedman, 2001). Doing so in this way intends to help stakeholders deal with complex issues through surfacing and addressing the multiple qualitative perspectives on an issue.

Analysis of content through map structure

The previous two points have addressed the relevance of the SODA methodology, in terms of map structure and content, to developing programme theory. Drawing these two ideas together in the SODA methodology is the analysis of map structure. Analysing the emergent structure of the maps helps to get a hold on ‘messy’ or complex issues through looking at structural features of the map (Eden, Ackermann and Cropper, 1992). Analysing the structural features of the map differs from analysing say pieces of qualitative text where one might look for themes or repeated ideas. With SODA maps, analysis of the emergent structure means that the way in which participants express their views on a problem area is translated in the map structure and subsequently provides the means for analysis. Analysis can be done by hand, but the use of software, such as Decision Explorer, can facilitate detailed analysis of more complex maps. Analysis undertaken will ultimately depend upon the purpose of conducting mapping and inquiry.

The emergent structure of the map facilitates identification of several interesting features which are be used to explore problems in more detail. For example, by calculating the total number of arrows in and out of a construct, we can readily understand central constructs, i.e. those ideas which impact on and are impacted by the most other constructs. Identifying those constructs which are central in the map in this way highlights the richness of meaning of particular constructs in the context of the whole map (Eden, Ackermann and Cropper, 1992). A further example of analysis which can be conducted on a map is based on its visual structure i.e. is it short and flat or long and narrow? A long and narrow map would indicate perhaps detailed argument without considering alternative perspectives on the problem (Eden, Ackermann and Cropper, 1992). A flat map might indicate that there is a wide variety of alternative definitions of the problem with little depth of understanding elicited about any one of them (Eden, Ackermann and Cropper, 1992).

In the case of developing programme theory, the analyses which the SODA methodology facilitates are helpful in terms of exploring mechanisms contributing to change, key programme components which are highly influential, and factors influencing desired outcomes, all of which are important aspects of developing and using a programme theory in the context of programme evaluation.

Mapping processes

As this study aims to understand the added value of *developing* programme theory for those practising evaluation in the context of third sector organisations, I argue that the SODA methodology is particularly useful as there is a heavy focus on stakeholder involvement (important also in the programme theory and evaluation literature) and on the process of building maps. Rather than being overly prescriptive, such as many approaches to using programme theory, the SODA methodology provides a set of tools to

deal with managing complexity and ‘messy problems’, that facilitate stakeholder involvement, and develop ownership of problem areas or issues. This is achieved through facilitating reflective negotiation and discussion about the nature of a problem in a manner that is also analytic through the analysis of map structure (Eden, 1988). Therefore, the flexibility of the methodology is useful in this setting as it allows the approach adopted to fit the programme evaluation practice circumstances, context, and challenges. There are different ways in which the SODA methodology can be applied and this affects the ways in which the maps are created as well as how they are analysed and used. There are two primary ways in which the SODA methodology can be employed: the cognitive mapping approach and the Oval Mapping Technique.

What follows are generic descriptions of the mapping process following in each of the organisations in which I developed programme theory. However, given the nature of this action research study, I wish to comment in more detail on the respective organisations, how the decision to develop programme theory was taken, and how each mapping process was implemented in Chapter 5.

Mapping process in organisation 1: cognitive mapping & facilitated discussion

Cognitive mapping aims to map an individual’s thinking about a problem or issue (Eden, 2004). Usually, cognitive maps are based on semi-structured interviews with stakeholders. In the interview setting, maps can be created in real-time. This is particularly advantageous as it allows the interviewer to clarify the meanings the interviewee attaches to certain constructs as well as to probe more detail from the interviewee (Eden, 1988) on interesting issues and to expand on and explain the meaning they attach to certain ideas. Alternatively, maps can be created in retrospect using the interview audio and/or interview transcript, although this approach leaves less opportunity to clarify and probe.

Importantly, and with respect to the action research methodology adopted for this study, the process of conducting interviews is said to be 'cathartic' (Eden, 1992) as it enables the individual to reflect, and in doing so change their way of thinking (Eden, 1992). As mentioned previously, individual maps can, if appropriate, be merged to act as a tool to facilitate decision-making and negotiation at the group level (Eden, 1992). Otherwise, the analysis of the structure of individual maps can be used as a tool for facilitated discussion when the merging of maps is not feasible and/or appropriate.

Mapping process in organisation 2: The Oval Mapping Technique

The Oval Mapping Technique (OMT) is another way to apply the SODA methodology. It is a tool to promote emergent strategizing in a group setting. In a group setting, OMT aims to develop ideas of 'what we know, what we do, [*and*] the way we do it', and captures this thinking as a structured map, or cause map (Eden and Ackermann, 1998, p.304), similar to that which I described in the section on map structure. However, as a group-based tool, OMT is described as 'quick and dirty' as it allows a group to surface their understanding of a problem together relatively quickly yet in a structured way (Eden and Ackermann, 2001). The map produced at the end of the workshop is a tangible outcome of the workshop. However it is in both the process of developing the map in the group setting as well as its resulting content that are *'powerful facilitative devices intended to help manage political feasibility … [through enabling] the team to arrive at something approaching consensus and both emotional and cognitive commitment to action'* (Eden and Ackermann, 2001, p.21).

The Oval Mapping Technique normally involves a workshop setting with around 6-12 participants. Using a rough agenda (Eden and Ackermann, 2001), workshops begin by setting the scene for the problem or issue to be explored. A period of brainstorming produces ideas transcribed on ovals that are stuck to a wall, followed by sessions where ideas are structured and iteratively

revised, considering which are goals/objectives, issues, options and constraints. In this way, there is a cycle of problem construction, making sense, defining the problem and ascertaining a portfolio of solutions (Eden, 1988). The iterative nature of the workshop and the discussion it facilitates are important parts of the process. The group-based process, as a result, aims to facilitate joint understanding, reflection, negotiation, and importantly commitment to action.

3.5.2.7 Data collection and analysis

An important element of ensuring the rigour of action research studies and the validity of findings is in documenting the data collection and analysis procedures through which the outcomes of the research can be drawn rather than simply basing research findings on researcher intuition (Eden and Huxham, 2002; Herr and Anderson, 2015). As per the nature of action research, data collection and analysis was an iterative, simultaneous, and emergent process. There were three key sources of data which I drew upon to generate learning from this study. Data collection focussed on collecting observations of practitioners reflecting-in-action, and in particular their reflections on their own programmes and their programme evaluation practice. The sources of data collection aimed to triangulate emergent learning based on my own and practitioners' reflections on the process of developing programme theory and its emergent content. The key sources of data collection are described below.

i. Observations of conversations and discussions by participants reflecting on the development of programme theory and its emerging content. These conversations and discussions occurred during the facilitated discussion and the workshop, but also in meetings in the lead up to the facilitated discussion and workshop. Observations from meetings were recorded in note form and relayed to those present in the form of meeting notes to facilitate a process of member-checking

where participants had the opportunity to question the content of my notes and reflections.

Observations from the facilitated discussion and workshop were recorded in note form, both during and after each event. These observations were noted as and when possible by the researcher during the events, but were also supported by another PhD student, who participated in both the facilitated discussion and workshop in each organisation. The additional PhD student acted as a 'critical friend' (Herr and Anderson, 2015) who shared her own observations on the process. Together we were able to discuss our observations which was useful for me given that I was focussed on facilitating each event and as such often missed certain observations or had forgotten them by the end of each event. In hindsight, I think it could have been beneficial to have audio recordings of the facilitated discussion and workshop; however, at the outset I thought this might have affected participants' willingness to reflect openly and honestly with each other and with me.

ii. I made personal reflections after the facilitated discussion and workshop. I recorded these reflections in note form. Primarily I reflected on the process of developing and discussing the programme theory and what this meant for the participants, reflecting again on my own observations of the discussions had.

I also reflected upon my own position as a researcher in each instance. I was able to clarify the emerging role I took in each organisation which I concluded to be more of a facilitative role than a complete 'outsider' role as in more traditional approaches to conducting research. I considered my facilitative role to align with what Herr and Anderson

refer to as an 'outsider in collaboration with insiders'[20] (2015) . I was able to reflect on how the role I adopted may have influenced the outcomes of the research. For example, I was able to support the organisations through the development of and interpretation of the programme theory. Without this support, it may have been the case that their ability to develop and interpret programme theory would have been lessened and as such they might have derived less value from the process. I consider the issue of reflexivity (my role as a researcher) in Chapter 5 which presents the findings of the action research study.

iii. Reflections from participants, during and after the action research intervention, were recorded in response to feedback questions. This was not in the format of a formal questionnaire, but collected primarily via e-mail after the facilitated discussion and workshop had taken place. Verbal reflections on these questions were also recorded in note form during each event.

During and after each event, I asked participants to reflect verbally, or in writing, on several aspects of the event: (1) reflecting on the value of the content of the maps (i.e. the programme theory) including things that were surprising, areas for further exploration, and potential uses of the information; (2) reflecting on the workshop/facilitated discussion and the process of developing programme theory in this setting; (3) changes in thinking as a result of the content of maps and process of workshop/facilitated discussion (in terms of both the programme and its evaluation); and (4) feedback on the format of the workshop/facilitated discussion and my role as facilitator.

[20] This refers to the fact that the research was being conducted primarily for my purposes (i.e. my PhD) and so I was the 'outsider' working with 'insiders' (participants within the organisation) to address some issue of concern to them.

3.5.3 Theoretical integration

An important aspect of employing a multi-methodological framework in in allowing for theoretical integration through the process of triangulation of findings (Moran-Ellis et al., 2006). In this study, as stated previously, I opt for a process of integration which uses separate methods, separate analyses of data, and presents separate findings. Integration then occurs at the very end, through a process of 'theoretical integration' (Moran-Ellis et al., 2006). I chose this method of integration due to the fact that the separate studies address different research questions relating to the research problem, and yield their own findings due to the descriptive and prescriptive nature of each research question.

Given that there is little conceptual understanding of the role of programme theory in small TSOs in the literature, the respective analyses and findings of each study are primarily inductive in nature. However in conducting the process of theoretical integration, which I present as the discussion section (Chapter 6) of this thesis, I opt to introduce new literature to generate insights by suggesting potential explanations for the emergent learning (Simpson, 2018). This way, the 'warranted assertions' I generate, based on the respective empirical studies, are explored using literature that helps interpret the findings in a more explanatory way. The implications of conducting integration in this way, provides the beginnings of a conceptual framework for future studies seeking to explore, in more depth, and with more novel insight, the use of programme theory in small TSOs.

3.6 Chapter summary

This chapter details and justifies the methodological framework adopted in this thesis. In adopting a pragmatic approach to empirical inquiry, I aim to generate 'warranted assertions' about the use of programme theory in programme evaluation practice in small TSOs through an empirical focus on the practitioner perspective. A multi-methodological framework enables me to

address the descriptive and prescriptive research questions respectively and with a dual focus on affording practitioners to reflect on their practices as well as reflect in practice.

4. CHAPTER 4 STUDY 1 FINDINGS

4.1 Introduction

This chapter presents the findings of the first empirical study conducted as part of this thesis. This study generates findings which relate to the descriptive question which this thesis addresses:

RQ1: What is the current practice with respect to the use of programme theory in programme evaluation practice in small TSOs?

a. What is the context of programme evaluation practice in small TSOs with respect to the capacity to do evaluation, evaluation activities, and the use of evaluation?
b. In what forms, does the use of programme theory manifest itself within the context of programme evaluation practice in small TSOs?

I present the findings of the qualitative study using the thematic analysis of the 23 semi-structured interviews with participants primarily from small third sector organisations but also participants representing three third sector funding bodies and two third sector evaluation support organisations.

4.1.1 Structure of findings

In the first half of this chapter, I present the findings of this study relating to programme evaluation practice in context. In this section, I comment specifically on the capacity to do evaluation (in terms of barriers and facilitators); evaluation activities (both 'systematic' and 'unsystematic' evaluation activities and the challenges to doing so); and the use of evaluation (in terms of the process of being involved in evaluation and the use of evaluation findings): these relate more broadly to programme evaluation practice in context. The findings reported in the first half of this chapter provide an understanding of the programme evaluation context which frames and provides insight for the findings presented in the latter half of this chapter on

the use of programme theory. The latter half of this chapter presents findings from 5 emergent themes on the use of programme theory in programme evaluation practice: the desire and need to use programme theory more explicitly; the implicit understanding of programme theory; the use of programme theory in 'unsystematic' evaluation activities; the use of generic theoretical frameworks; and the use of 'individual-level' programme theory. The order of the findings in the latter half of this chapter reflect, in my opinion, the order of importance of the use of programme theory in evaluation practice. Finally, I summarise these findings in terms of putting into context the current practice in the use of programme theory in evaluation practice in small TSOs. The findings presented are summarised in Table 6 for the readers reference.

Focus of empirical research		Theme(s)	Sub-theme(s)
[1] EVALUATION PRACTICE IN CONTEXT	Capacity to do evaluation	Barriers	Time, core funding, resources, structure, bureaucracy, funder expectations
		Facilitators	Understanding of change (small size, service delivery role, professional training, communication)
	Evaluation activities	'Systematic' vs 'unsystematic' evaluation activities	
		Practical challenges in conducting evaluation	Capacity for beneficiary participation, power and ethics in evaluation, purpose of evaluation
	Evaluation use	Benefits of being involved in the process of evaluation	Learning and reflecting, benefits for beneficiary involvement
		Use of evaluation findings	Use of 'systematic' and 'unsystematic' evaluation findings
[2] USE OF PROGRAMME THEORY IN EVALUATION PRACTICE		Desire and need to use programme theory more explicitly	
		Implicit understanding of programme theory	
		Use of programme theory in 'unsystematic' evaluation activities	
		Use of generic 'theoretical' frameworks	

	Use of ‘individual-level’ programme theory

Table 6 Summary of findings on the use of programme theory in the context of programme evaluation practice in small TSOs

4.2 Findings

4.2.1 Programme evaluation practice in context

Initially I present the findings of this study relating to the overall context of evaluative activity within this sample of small TSOs. This section presents findings relating to the capacity to do programme evaluation, evaluation activities, and evaluation use in small TSOs i.e. evaluation practice in context.

4.2.1.1 Capacity to do evaluation

In Chapter 2, I discuss how the capacity to do evaluation is defined as the skills, knowledge, and motivation to do evaluation along with the organisational support structures that facilitate evaluation. Given that this sample of participants from small TSOs is classed as ‘accidental’ evaluators, capacity to do evaluation is an important factor to consider in influencing the ability to do programme evaluation. In this section, I discuss the findings of this study in terms of the barriers to and facilitators of the capacity to do evaluation in this sample of small TSOs. Like other studies of evaluation practice in small TSOs, barriers include time, core funding for evaluation, resources, and structure of evaluation processes, as well as the perceived bureaucracy with respect to evaluation and funder expectations of evaluation activities. Facilitating the capacity to do evaluation includes the knowledge of change, as noted in small organisations that exhibit close contact with beneficiaries and staff, as well as field-specific professional training.

1. *Barriers*

Similar to other studies of evaluation activity in third sector organisations, participants (third sector organisations, funders, and support organisations)

indicated that time, core funding for evaluation, and resources were broadly considered to affect an organisation's ability to conduct evaluative activity. Reflecting on time, funding, and resource constraints, one participant mused on the gaps between funding for service delivery and funding for the core costs of an organisation. She perceived more stability in core funding, to support more effective evaluation within the organisation, but that core funding was more difficult to fundraise for. She told me, *"there's not enough money invested into the core parts of the organisation … it creates a gap we have to fundraise for … it's much easier to fundraise for outcome delivery"* [Org J].

Particularly in the case where there have been funding cuts, for example to administrative capacity, programme managers increasingly find themselves working in service delivery and taking on multiple roles within the organisation. This means, in many cases, that evaluation activity takes lower priority within small TSOs. One project co-ordinator reflected on the lower priority given to evaluation due to the continual need to be providing services to young people. She said, *"we're not very good at blowing our own trumpet about what we do because we just get on with doing it and when we're finished something, we're away doing it again or we've got another group of young people"* [Org D].

On a practical level, some participants considered the barriers in terms of time, core funding, and resources meant a subsequent lack of structure to their evaluation processes. Several participants identified this as having a constraining effect on their ability to conduct evaluation, resulting in implications for the perceived value of evaluation activities. In these cases, a lack of structure leads to more 'ad hoc' approaches to formal evaluation activities, and difficulty in pulling large amounts of data together. In terms of a lack of structure, a community representative from one organisation noted the inability to evaluate the impact of the programme beyond the day-to-day operations. She said, *"We don't know how to figure out how well we're really doing. We know how we monitor and review what we're doing on the ground*

face-to-face, but we don't know how to monitor the rest of what's happening, the ripple effect" [Org M]. Relatedly, a board member in another organisation reflected on the implications of a lack of structure in terms of failing to capture potentially useful data. She reported that, *"missing out on some of that data is missing out on proving we've done what we said we would … and there is something to that, you can say, 'we did this, and we did that' but if you have nothing to back it up with then that's a problem"* [Org G]. Similarly, a staff member from another organisation reflected that a lack of proper and purposeful structure to evaluation processes meant *"that the format of our evaluations are quite bland and so, although you know the aim of the group, due to the format of the evaluation, you don't know how to justify things properly so you speak kind of blandly about things"* [Org A].

Some of the participants commented also on the bureaucracy of evaluation and the impact that bureaucracy has on their ability and importantly willingness to do evaluation. One participant stated that some employees in her organisation had made a move to practice in the third sector in order to escape the bureaucratic systems of the public sector, allowing her to focus on service delivery and improving the lives of those with whom organisation works. She said "*the thing is a lot of those people who used to be youth workers or were council employees in some sort of support mechanism, the reason they left is because they did too much paperwork. You find them in these organisations now, which feels much more liberating to them, and as soon as you show them a form it's like 'oh no here we go, this is bureaucracy creeping in and I just want to get my good work done'"* [Org G].

Participants from funding organisations noted that they needed to be clearer in what they required in terms of evaluation, whilst remaining sufficiently flexible to allow third sector organisations to conduct evaluation in ways that fit the third sector organisations' needs and capabilities. Participants from funding organisations acknowledged the need to re-dress this balance in terms

of simplifying the reporting of programme evaluation, and thus ensuring the programme evaluation processes of TSOs can cater to the needs of both TSOs, and funding bodies, whilst maintaining proportionality. One representative from a funding organisation told me, "*One of the things we're looking at in the training is to try to ensure that they recognise that we're trying to keep things proportionate for them but we need to better articulate what that means and how that's reflected …we need to get it in a language that doesn't put people off doing it"* [Fund 3]. Despite this, the power imbalances in funding relationships are acknowledged in funding organisations. The same participant told me about trying to redress the balance of power by reinforcing the idea that without the third sector organisations, funders can do nothing with the money they hold. She told me, "*money is the power, but we could be sitting on loads of money but we wouldn't be doing anything with it … if it wasn't for them. We've got to constantly reinforce that message in terms of our behaviour, but also in terms of the information we've got, the language that we use, everything"* [Fund 3].

2. *Facilitators*

I found that there are some factors that seem to facilitate the capacity to do evaluation. Indeed, many of those conducting evaluation in such small TSOs also had a front-line service delivery role to play. Whilst this is often perceived as a challenge, due to restricting time available for tasks like evaluation, in some cases, participants commented on the fact that this meant they knew exactly what was happening as a result of their daily contact with the programme, and programme beneficiaries. One project manager told me that, *"it means I'm in touch with all those things going on. There's nothing that happens in the project that I don't know about… That's probably the luxury of being a small organisation that has a lot of clout"* [Org E]. In this instance, the participant acknowledged that his knowledge of everything that was going on was advantageous in terms of evaluation as he felt he was more 'in touch', or aware of, information that could be used in evaluation activities.

Similarly, the closeness to the beneficiary was considered advantageous for conducting evaluation, given that knowledge of change was observed, on a day-to-day basis in many cases, by those conducting evaluation. In one case, a participant reflected on the value of being back on the 'front-line'. She was required to move back into service delivery after funding had been cut and told me that, *"I think because I'm forced back on the front line, I'm in amongst it myself…and you actually start to get a feel for what's working and what's not working … - you just get a feel"* [Org A].

Further to the front-line nature of the role of many of those practising evaluation in small TSOs, and the subsequent closeness to the beneficiary, participants commented on the fact that the practical focus of their professional training, in for example youth work or social work, facilitated strong understanding of beneficiary need, and service delivery to improve outcomes for beneficiaries. One participant reflected in particular about his staffs' ability to bring expertise from other or previous jobs to improve the services they offered in the project. He said, *"they're very good at their jobs…because of the fact that a lot of them do or have worked in the social care aspect as well as the fact that they're bringing other traits from other jobs to our team"* [Org C]. Talking about the experience of social work and personal challenges of the founder of another third sector organisation, one board member also told me, "*I think she would suggest that everything she does is developed from a combination of professional experience and personal experience as well as having tried 'trial and error' in developing things"* [Org G]. The nature of the expertise from professional practices and experience in such cases facilitated evaluation activity, particularly on a day-to-day and more informal basis, primarily through being reflective on past experiences and open to learning from peers.

However, whilst resource was considered a constraint in terms of factors such as expertise, the small size of these organisations facilitates formal and informal communication channels between members of staff, leading to better evaluation, albeit on a more informal and less systematic basis. The opportunities for reflection and peer learning were in many cases reported to be facilitated by structures within the organisation that gave staff the opportunity to come together regularly to share experiences. For one participant, this came in the form of an overnight residential, where staff from similar projects gathered to share learning. She told me that, *"it's about all the projects coming together and sharing their challenges, their successes, what's worked well … in terms of our own personal development, that can be really valuable"* [Org F]. In terms of evaluation, she said that in coming together in this way she could *"hear someone talking about an evaluation that they've done and think that sounds really good and that would work really well with our group if I maybe adapt it a little"* [Org F]. In other cases, staff engaged in such reflection and peer learning on a day-to-day basis, or in weekly team meetings, for example.

Relatedly, communication is facilitated, in some cases, by trust as evidenced in perceived open cultures and ethos around sharing of experiences and practices of what is working, and how. This is perhaps further influenced by the emphasis in professional practices, such as youth and social work, on being reflective. A CEO acknowledged the need for her to allow staff the time and space to get evaluation activity 'right'. Reflecting on a recent overhaul of evaluation processes to integrate participation into evaluation, she said, *"how I've done it is really fighting against my impatience and allowing the staff to do it at their own and speed …it's got more of a chance of success of they are brought into it and are part of the design of it"* [Org J]. Moreover, there were other examples where participants observed that there was an open and honest culture around evaluation, facilitated by the collaboration and small size of the organisation. For example, one participant told me, *"we've got a senior*

management team that works closely with staff…therefore there is a lot of communication" [Org K].

4.2.1.2 Evaluation activities

In what follows I present the evaluation activities reported to be practised in this sample of small TSOs. In Chapter 2, I defined formal evaluation activities as those which involve systematic inquiry, at any stage in the programme life cycle, to describe or explain the programme in terms of its implementation, effects, and social implications; such systematic inquiry can be conducted using a variety of methods and approaches (Mark, Henry and Julnes, 2000, p.3). I also highlighted that 'thinking evaluatively' forms an important aspect of formal (systematic) evaluation activities in supporting the ability of evaluation practitioners to plan, design, and carry out formal evaluation activities. Further to this I present findings which reflect the challenges in conducting evaluation activity as experienced by those practising evaluation in small TSOs.

The key findings presented here reflect that small TSOs conduct two 'types' of evaluation activities, those which are 'systematic' and those which are 'unsystematic'. I do not wish to make a good vs bad distinction by using 'systematic' and 'unsystematic'. Rather I want to reflect that evaluation practice in this sample seem to be carried out in a 'systematic' way in terms of reporting to funders where specific methods and approaches are used to collect and analyse data. On the other hand, by 'unsystematic' I seek to reflect evaluation activities that happen more informally, on a day-to-day basis. These activities do not necessarily use a specific method or approach, rather they illustrate the on-going reflective practice of practitioners which informs their everyday work rather than information which is systematically gathered and analysed (Mark, Henry and Julnes, 2000).

'Systematic' evaluation activities

Similar to other studies of evaluation practice in third sector organisations, much of what constituted 'systematic' evaluation activities focussed on reporting to funders. In many cases, that means reporting to multiple funders throughout the year. On participant told me, *"I suppose evaluation is embedded in the organisation because I have to report to every funder and we have 20 plus funders so we have to do an evaluation report for each"* [Org A]. I found evidence that much of what was considered to be 'systematic' evaluation was that which is explicitly conducted and written up was for the purposes of writing reports to funders. Programme evaluation in this regard seems to be perceived as an explicit activity conducted at a specific point in time i.e. annually. It seems as though much evaluation activity which was considered for input to reports, in a 'systematic' sense, was collected using a combination of methods including gathering qualitative feedback, using participant surveys, developing case studies, developing and using activity-based evaluation formats, and analysing output and monitoring data. As one respondent described it, *"we do quite a lot of things, sort of a mosaic approach to it"* [Org C].

'Unsystematic' evaluation activities

In conducting conversations with participants regarding evaluation activities, and, in particular, in asking about how services were designed and monitored, I found that there was much evaluation activity considered throughout the lifecycle of the programme on a more informal basis. This kind of evaluation activity was not necessarily systematic in nature, in terms of not being planned and conducted using specific methods for data collection and analysis, rather, it was part of the day-to-day operations of small TSOs. In addition, participants did not necessarily consider this 'unsystematic' evaluation activity to be 'programme evaluation' at all. Nevertheless, it reflects a less systematic approach to developing, adapting, and improving their services and programmes. What I considered to be 'unsystematic' evaluation activity varied

across organisations, but broadly involved programme staff members coming together, whether on an ad-hoc basis or during arranged meetings, to discuss the programme, how it is working, what it is achieving, and what they could do better for programme beneficiaries.

In one organisation, 'unsystematic' evaluation activity took the form of a weekly staff meeting. She told me that, *"we have a weekly staff meeting and that's about sharing feedback about what's going on as well as identifying the ongoing issues or problems…we try to resolve problems, we note successes and we see what strategies have worked" [*Org C]. Doing so, her colleague told me, facilitates a lot of peer learning through sharing of best practices [Org C]. Other organisations reflected on, what was referred to as informal evaluation activity, as something that happens on a continual basis. One participant stated that *"we are in a constant state of feedback and evaluation and it can be from one-to-one sit-down chats to talking to people on the street, to talking to communities. It's that talking and feedback …* [that allows you to*] gauge where you're at and what impact you're actually having"* [Org E].

'Systematic' vs 'unsystematic' evaluation activities

Reflective practice, informal feedback and communication from and between beneficiaries and staff, and day-to-day observations were considered on a more informal and less systematic basis, rather than being captured by specific methods of data collection/analysis before collation in a formal write-up or report. This seems to be the case because often 'unsystematic' evaluation activity, undertaken by participants in small TSOs, is not considered to be programme evaluation.

However, some participants noted that the kind of information, produced by such day-to-day unsystematic evaluation activities, could be better collected in 'systematic' evaluation activities through more structured approaches to

collecting and analysing data. One staff member reflected on the fact that staff were aware of the achievements of the programme but did not capture it in more systematic evaluation because of the nature of the information. She said, *"I think we know that we're doing a good job but I think like, in terms of how it's recorded, I think we could be stronger on that because then obviously it's better for [PROJMANAGER] to then collate all the information and write a report on it. It is difficult to capture in evaluation the fact that a lot of the success of the work comes down to individual relationships with young people"* [Org A]. Similarly, another participant reflected on the perception that some of the information or knowledge that staff hold is not eligible for more 'systematic' evaluation efforts stating that, *"there can be a feeling that if it's not a number then it's not data or it's not valid for inclusion"* [Org G].

One suggested reason for the lack of inclusion of 'unsystematic' evaluation activities in 'systematic' evaluation activities are the professional perspectives, i.e. social work or youth work, held by those conducting evaluation in small TSOs. In such cases, strongly held values of improving the lives of beneficiaries is reported to be a primary driver of the work that they do and is embedded within third sector approaches to working. One participant reflected on the important role played by the previous experience of the project co-ordinator in her organisation. She said, *"I think* [project co-ordinator] *would suggest that everything she does is developed from a combination of professional experience and personal experience…she was a social worker before she started this so she's got a certain level of experience there and coming from the background that she has had, she has seen a lot of the ways that have helped in her own recovery from the challenges she has faced"* [Org G]. Here, the participant was reflecting on how both professional and personal values can help to shape the way programmes and projects emerge, change, and improve. In such cases, it might be that such professional and personal values are difficult to 'systematically' disentangle and evaluate. Moreover, the professional backgrounds, i.e. social work or youth work, of many practitioners embeds the notions of reflective practice and improvement as core

components of the job and as such perhaps encourages continual cycles of reflection and action.

Some participants also commented that the lack of structure in terms of 'systematic' evaluation activities meant that the valuable understanding (from 'unsystematic' evaluation) was not often captured in a systematic sense. In one organisation, a staff member reflected on how the learning from individual sessions was not translated into formal (systematic) evaluation due to a lack of appropriate methods for recording such data. She noted the knock-on effect this had on formal evaluation activities. She said, *"I think if it was recorded better in our evaluation, it would make it easier for us to reflect … and we know it has to be addressed because it makes it difficult for* [project manager] *to go back and give a report"* [Org A]. Another participant similarly reflected on the lack of structured format to capture feedback on a daily basis. She told me, *"we've never actually given out a feedback form for parents, but verbally we hear feedback … but maybe we should have some sort of template to try to capture what the community actually think about the project"* [Org B].

The role of funders was also noted as important here. 'Unsystematic' evaluation practices, with a focus on continually improving programmes and services, means that whilst the day-to-day operations are adapting and improving, such information is not necessarily reflected in systematic evaluation activities, despite this learning being of interest to funders. One participant, representing a funding body, reflected on their organisation's interest in understanding how small TSOs have improved services over time. She said, *"I suppose we talk a lot about this whole thing about 'improving versus proving' … I think in the past there has been too much emphasis on prove rather than improve. I would like to see that it's* [evaluation] *has been useful for purposes of improving"* [Fund 1]. She went on to describe their most recent reporting requirements which *"are not about outcomes specifically but they ask about the difference that has been made, but also about what they*

[small TSOs] *have learned and how that can be used".* Other participants from funding bodies noted however that some small TSOs might be afraid to share some of the learning around improving as it might indicate weaknesses. One participant stated that, *"some TSOs are much less comfortable with sharing the challenges because the funding climate has become so competitive that they feel if you show any signs of weakness, either funders will say they can't fund them or other organisation will say they can do things better than them"* [Fund 2].

1. *Challenges in undertaking evaluation activity*

Challenges in actually undertaking evaluation activities, particularly more 'systematic' evaluation activities, were noted by many participants, often from the perspective of the beneficiary. Some of these challenges included: the capacity of programme participants to partake in evaluation (e.g. literacy issues); power issues with respect to who and what is being evaluated; the ethics of conducting evaluation more generally; and the challenges posed by funding requirements in which case there was scepticism surrounding how reports for funders were actually being used to generate action and change.

i. Capacity of beneficiaries to partake in evaluation

Some participants, particularly those working with vulnerable groups, commented on challenges in engaging participants in evaluation. Participants reported issues of confidence, and feeling intimidated by an evaluation form that can act as barriers to beneficiary participation [Org H]. One participant reflected specifically on the issues with literacy in the groups with whom the project was working. Whilst case studies were a valuable method for collecting and presenting data for 'systematic' evaluation activities, she said, *"we try as much as possible to get the case studies in people's own words. Obviously, there are a lot of literacy issues in the people we work with and that again can just be me scribing what someone is saying to me. We do have recording devices that I'll sometimes use"* [Org F].

Some interview participants reflected on strategies to overcome challenges in engaging vulnerable groups in evaluation activities. One participant recognised that the ability to engage the young people with whom the project was working, was limited due to their disabilities. She commented on actively seeking out new ways to engage with young people to overcome this challenge: *"we do consult a little bit with the young people but it's very much at a smiley face level … and what I'm trying to do is to give the young people a chance to take photographs of the sessions so we're getting their point of view … and* [through that] *we're hoping to get a bit more information about how they feel … we might get an idea of what's important"* [Org C]. Another participant acknowledged that they often found young people filling out evaluations in a rushed manner to 'just get them done'. However, in realising this, he reflected on their strategy to overcome this challenge to generate more useful information, for evaluation purposes but also for the benefit of the young person. He told me, *"we help them with that, because young people tend to just tick things for the sake of it. So, we're trying to sit down with them and think about what they're writing and explain themselves, what they're trying to do, and what they think…they're actually learning to express themselves"* [Org E].

ii. Power and ethics in evaluation

Another challenge realised by participants was the power dynamics and ethics of conducting evaluation, particularly with vulnerable groups. Specifically, this related to issues of determining whether outcomes had been achieved for programme participants and using programme participants as sources of information. One participant, from a TSO, stated that conducting evaluation felt like you were making judgements on their lives. He told me, "*there's that sense that you are evaluating somebody else's life, and I think, ethically, there's a bit of tension with you seeing whether somebody's life is successful or not"* [Org H]. Relatedly, another participant reflected a sense of protectiveness over

conducting evaluation *on* an individual. In acknowledging that not all the information they hold should be used for 'systematic' evaluation activities, she said, *"I think there's a lot of work that goes on that could be used as evaluation information but there is also a protectiveness over the clients, over exploiting the clients and using clients for information in that way … it's a difficult line I think"* [Org G].

Overall, many participants were conscious of maintaining balanced power dynamics and treating often vulnerable groups in ethical ways so as not to exploit them for evaluation purposes. Reflecting on this, one participant told me about ensuring participants felt like they were not being exploited by using more relationship-based approaches to evaluation. He said, *"what the process of the informal* [unsystematic] *stuff does is it makes it more personal…it takes away that intimidation factor that people aren't good with: literacy or just unconfident in reporting on things: that can be a huge barrier, especially if you just give them a form. If you remove that and have a conversation, you make it about them…that is a powerful thing for the job that we do, it's 100% based on relationships and trust"* [Org H]

iii. Funder use of evaluation reports

A further ethical challenge experienced in small TSOs was related to the conflicting purposes of conducting evaluation for funders. In particular, in contrast with the unsystematic evaluation activity being conducted on a day-to-day basis, one of the challenges experienced within small TSOs was a perceived feeling that systematic evaluation activities, often conducted for funders, had little purpose because they were not used to contribute to any meaningful change, for the funding organisations or for the third sector more widely. One participant reflected on ensuring evaluation activities remained reflective and purposeful within his organisation. He said, *"in terms of being more formalised, it* [evaluation] *has to have a point, a purpose, otherwise it is*

just a task. I think, whenever you start doing tasks, that's when things are professionalised, because tasks have to be done in particular ways" [Org H].

This was also reflected by participants from funding organisations who, in all cases, commented on the fact that as funders, they needed to be clearer on how the information provided in evaluation reports was being used to generate positive change and action. Some participants from funding organisations acknowledged that, whilst they were trying to improve how they used the information from reports from small TSOS, that they were not currently making effective use of such information. Some participants from funding bodies noted recent efforts to review processes to improve the use of information from evaluation. One participant said, *"I've being doing a lot of work looking at how we can use the evidence to better inform out own work to inform our funding decisions … and how we can gather useful learning and share that with the sector…part of that that is to go through* [the evaluation reporting mechanism] *and make sure that for every question we asked, we knew how we were going to use that information"* [Fund 2].

4.2.1.3 Evaluation use

In Chapter 2, I discuss how the capacity to do evaluation, and the ability to 'think evaluatively', enables evaluators to better conduct evaluations that are both feasible and useful for evaluation stakeholders. I now discuss the findings in terms of how evaluation activities are actually used. I discuss findings in terms of both process use, and findings use, where process use refers to the benefits of being involved in the evaluation process and findings use refers to the use of the findings of evaluation activities.

1. *The process of evaluation*

Some participants mentioned that the actual process of conducting evaluation, and, in particular, the 'unsystematic' evaluation activities I discussed in the

preceding sections, facilitated a sense of continuous learning and reflection within the organisation. Participants in organisations that had the opportunity to continuously learn and reflect on what they were doing, and why, experienced a variety of benefits.

One participant commented on the use of peer learning groups to facilitate the sharing of best practices, where he told me, *"we see a lot of peer learning through the staff groups … because they pass on if somebody has responded well to something or the young people responded very badly to something or if there's ways they have to engage a lot better with them then try to share the best practice and some of the practices"* [Org C]. Through these more unsystematic processes, examples such as the peer learning in Org C, reflect benefits which can be used beyond any specific evaluation activity. In this case, the staff are embedding the notion of sharing and using 'best practices' to inform their work more generally. Moreover, by having the opportunity to continuously reflect, participants commented on feeling more motivated and valued in their roles through being able to continuously reflect on programme achievements and the resulting positive changes occurring in peoples' lives.

Having the opportunity to continuously reflect also enabled the participants to feel that they could be more responsive to need, often through being more innovative and risk-taking in their practice. In some sense, the process of continuous reflection seemed to give some participants confidence to be innovative in their practice, based on their knowledge of the programme and its beneficiaries, and thus to try new things. Many participants for example reflected that the 'trial and error' nature of their work allows them to be responsive and innovative. One participant, reflecting on the role of evaluation, said that as a result, *"I think we're more in-tune, we're more able to respond. I think we're more thought-out and I think it forces you out of your comfort zone because it encourages you to respond to things that are quite scary to respond to"* [Org A]. In this sense, the participant felt that the evaluation of programmes

and projects facilitated a sense of confidence in implementing effective strategies to addressing specific problems.

Some of the participants commented on the integration of evaluation activity with service delivery. This meant that, in several cases, the process of doing evaluation was also considered to be beneficial to the programme beneficiaries. Integrating evaluative activity within service delivery was done in various ways, including activity-based evaluation, i.e. creative activities, and using components of service delivery as a means of also evaluating change. In this way the process of evaluation was considered empowering to beneficiaries but also beneficial in terms of generating insights about the programme and changes in outcomes as a result of the programme. One participant commented on their organisation's recent shift to more participatory approaches to evaluation stating that, *"whenever I thought about how much better it is to involve people in the process, that is the point you start seeing real results when people are involved throughout, rather than just being evaluated at the end"* [Org H]. He went on to describe the integrated process of evaluation in service delivery in *"having deep, meaningful evaluative conversations with people about what has changed for them also allows us to get to know them a bit better and it helps them trust us a bit more because you can actually see them reflect … these are skills and tools that will hopefully benefit them beyond just evaluating them"* [Org H].

2. Use of the findings of evaluation

'Systematic' vs 'unsystematic' evaluation activity findings

Similar to other studies of evaluation practice in third sector organisations, the findings of what was considered to be more 'systematic' evaluation activities were often included in reports to funders and board members. Moreover 'systematic' evaluation findings were included in communications and promotional materials which are used, internally and externally, to disseminate programme information to current and potential programme participants and

the wider communities in which programmes are based. On the other hand, it seems that the information obtained mainly from 'unsystematic' evaluation activity is used to plan and design, adapt, and improve the services and activities offered by programmes on a day-to-day basis (I presented this finding in the sections on evaluation activities). One participant reflected on the use of feedback mechanisms from young people her project works with. She said, *"if the feedback is coming back that a particular activity is not being well received by young people then we would obviously aim to change that piece of work or adapt it or just completely lose it altogether"* [Org B]

As with being involved in the process of evaluation, the findings of evaluation activities in terms of hearing and recording feedback from beneficiaries, humbled and motivated some participants. One participant told me about how staff use the findings of evaluation, perceiving the findings as often motivating as it gives them a chance to see 'on paper' the changes they are making. She told me, *"there's encouragement for staff because they can see the change that they're instigating with young women. They can see the results and the progression that's happening with the evaluation tool. They, in turn, are motivated by the change that we're seeing"* [Org J]

4.2.1.4 Summary of programme evaluation practice in context

The preceding sections aim to give more detail on the context of programme evaluation practice in small TSOs. I find that whilst the issues of funding, time, resources, lack of structured evaluation processes, and bureaucracy were considered to be challenges to conducting programme evaluation, the closeness to the beneficiary and field-specific professional training were considered to facilitate programme evaluation through enabling better knowledge of change for beneficiaries. Whilst there was some evidence of formal (systematic) evaluation activities in small TSOS, there seems to be a lot of evaluation activity occurring which can be considered as 'unsystematic' which happens more informally and on a day-to-day basis. Certainly, being

involved in more 'unsystematic' evaluation appeared to benefit staff both through the process of being involved as well as facilitating the improvement of programmes and services through sharing of experiences. More systematic programme evaluation efforts were reported to be used for reporting to funders and for communication purposes. Whilst these findings are not wholly different from previous studies, they provide a good understanding of the context of programme evaluation in this sample, supporting and contextualising our understanding of the following findings on the use of programme theory in current programme evaluation practice.

4.2.2 The use of programme theory in programme evaluation practice

The aim of this study is to explore the current use of programme theory within the context of programme evaluation activity in small TSOs and to investigate the potential contextual influences on how to use programme theory. In the following paragraphs I present the findings relating to the current use of programme theory. In many cases, participants were unfamiliar with evaluation terminology such as 'programme theory' and so my presentation of these results is dependent on an interpretation of the use of programme theory in evaluation practice. As such, in the following sections I explain why examples given are reflective of the use of programme theory. Specifically, I consider the desire and need to use programme theory, the understanding of programme theory, the use of programme theory in unsystematic programme evaluation activity, the use of general theoretical frameworks of change to guide practice, and the use of 'individual' level programme theory.

4.2.2.1 Desire and need to use programme theory more explicitly

The first key finding I present with respect to the use of programme theory is the desire and need to use programme theory more explicitly in systematic evaluation activities. Importantly, this was not just a funder-imposed desire and need. Rather, there were also some suggestions from the TSO perspective

that reflected the desire to make more systematic use of programme theory in programme evaluation.

TSO perspective

Whilst none of the organisations commented on the explicit use of programme theory in more formal (systematic) programme evaluation activities, many expressed a desire and potential need to do so in the future. This desire and need was expressed for a variety of reasons including shifting the use of programme theory from the level of individual activities and sessions to the level of evaluating the whole programme; to manage the programme better as a whole; and to have a framework for formal evaluation that is integrated with the work of the programme, so as not to overburden staff.

In terms of better using programme theory to understand the programme as a whole, one participant told me that her organisation use a framework in every youth work session that asks them to define the activities, outcomes, and overall aim of that session, a similar idea to programme theory. However, she acknowledged that it might have more value if applied at the level of the project as a whole, to facilitate better recording of information at the project level. She told me, *"we are encouraged to use that framework with every session …I think if we applied that same tool to our evaluation it would make things better because everyone knows what our aims are and what are outcomes are, but I think if it was recorded better in our evaluation, it would make it easier for us to reflect"* [Org A]. She perceived that recording information in this way at the project level would facilitate better reflection on the project as a whole.

In a similar way, another participant reflected the need to think more about the overall goals of the programme and how these were being achieved at the level of the programme. She reflected on how the kind of understanding provided by better reflecting on the programme in this way would allow her, as

project manager, to better manage the programme. She reflected on the potential role of programme theory, *"to have the programme more structured …it's really to re-evaluate what we're doing, why we're doing it, who is responsible for doing it and who is accountable"* [Org B]. In this case, the participant reflected more on the potential role of programme theory in terms of programme management. In this sense, the understanding, at the level of the programme, concerned what was happening, why it was happening, and its overall goals could facilitate better programme management and improved structure in programme delivery.

Whilst some organisations reflected on the desire or need to use programme theory more effectively, others had concern over overburdening staff with additional programme evaluation requirements. However, some felt that using programme theory would provide a solution which was more aligned with the work that was happening on the ground. One participant reflected on some of the evaluation training materials provided by SUPORG1 which were aligned with programme theory in defining programme activities, outcomes, and aims. She said, *"it felt very common sense to me whereas some of the other approaches you hear about, I don't particularly like them…I'm much more inclined to go for something that is integrated into the work that is already happening and try not to overload"* [Org G]. In this example, the participant was reflecting on the fact that by defining programme activities, outcomes, and aims, in one framework for evaluation, staff might be able to better appreciate the relevance of an evaluation framed in a way which is closely aligned to how they deliver the programme, rather than some abstract evaluation framework.

In noting the potential value that the value programme theory can have for small TSOs, a participant from a third sector evaluation support organisation told me about a training workshop he runs, based on the principles of programme theory. Besides the value of using the programme theory framework in systematic evaluation activity, he thinks that there is potential

value from simply thinking through the idea of programme theory more explicitly. He told me that the exploration of how programme activities links to outcomes, *"the thinking through of that, it's not a form-filling exercise, it's a thinking exercise … it's the difference between a logic model and logic modelling…once people realise that it's not a box-ticking exercise, it's actually a thought-exercise about planning…I think they see that people would describe the benefit of the process. Then they have a usable tool, the finished product is a bonus"* [SupOrg2]. In this case, the participant reported that participants in his workshops tend to find the process of developing programme theory beneficial, although he could not comment on the extent to which these benefits were actually realised in the actual evaluation practices of the organisations he trains.

Funder perspective

In contrast to this desire and the need to use more programme theory-inspired approaches, organisations commented that this is not something that they were explicitly asked for by funders and thus did not feel the need to spend time and money to employ such approaches. By contrast, participants from funding organisations were very much interested in the principles of programme theory. In particular, they were interested in how such approaches could enhance their own learning about effective strategies to tackle social issues in order to both facilitate more strategic funding approaches and to share learning with the third sector more widely. Funders also noted a desire to shift towards a more learning-focussed approach to reporting rather than solely an outcomes-focussed approach. One participant told me, "*we're more a relationship-based funder, we are not as rigid because things change and the approach to outcomes that we used before was a little too inflexible I think…I suppose you would want the organisation to reflect that kind of learning culture to show that they're using their own evidence for purposes of improving, or whatever other purposes they want to use it for"* [Funder 1]. Whilst this quotation does not explicitly refer to programme theory, the

participant emphasised the use of evidence for purposes of improving, i.e. using what he knows about how the programme is working so as to improve the programme.

However, participants from funding organisations noted a specific challenge on how best to elicit this kind of information. In order to avoid becoming overly prescriptive about the kinds of information sought they had to be flexible in their approach to funding reporting requirements. In one case, the participant reflected on the change in their approach to reporting. She said, *"we've swung from being quite prescriptive to much more freedom…people appreciate that the requirements are less than they used to be. However, some groups, particularly less experienced groups, like the reassurance of a template because it's just clear about what our expectations are…obviously internally for us it makes it a lot easier if there is some kind of standardisation"* [Funder 1]. In this case, the challenges are in balancing the reporting formats between having some sort of standardisation, so that TSOs have something to work from, but also that the funding organisations can draw out learning from across reports.

Similarly, another participant from a funding organisation reflected on maintaining a balance between proportionality, i.e. the asking of information in a way which is amenable and accessible to small TSOs, and ensuring better quality reports through their evaluation reporting requirements. She commented on an ongoing review of their training and reporting procedures. She said, *"we're looking at our training to ensure that we keep things proportionate for them* [small TSOs] *but we need to better articulate what that means and how that is reflected and we need to get it in a language that doesn't put people off doing it…we are hoping that these will then be better reports, easier for you because you know what we want, easier for us because we've got the quality"* [Funder 3]. In this case, the participant acknowledged

the role that they, as the funding organisation, need to play in asking for and generating the programme theory related information that they want to gather.

4.2.2.2 Understanding of programme theory

From the participants in this sample, it was clear that there is a deep understanding of and willingness to *talk* about the programme theory, i.e. how the programme works and why it works, even if not using the 'programme theory' terminology. In many cases, participants talked about what it is that their programme or project was trying to achieve overall through describing general approaches to change contrasted with alternative approaches to change. In the following examples, participants described at a wider level how their programmes and projects aim to instigate change.

For example, one participant reflecting on her project's approach to tackling youth disengagement in the local community, through a youth project, criticised other approaches to change in contrast with that taken by her project. She said, referring to local government, *"I mean they would spend millions of pounds trying to stop them throwing stones at buses and smashing windows and I thought you need to stop doing that … it's about positive reinforcement here"* [Org D]. In this example, the participant contrasts two different ways, or mechanisms, to tackle youth disengagement, i.e. positive vs negative reinforcement. In this case, the idea is that this particular programme adopted the principles of positive reinforcement in designing programme activities.

Similarly, another participant spoke about how her approach to change in their programme contrasted with traditional approaches to dealing with trauma and traumatic experiences. She said, *"You're really trying to develop the skills that they have and help them engage…rather than fall back to the community where the trauma was perpetuated usually…we are trying to encourage independence in them …it's a social model rather than a medical model"* [Org

J]. The overall approach to change, adopted within this example, aims to promote social changes with respect to dealing with trauma, in contrast to a traditionally medical approach to change. In this way, the programme supported young women to rebuild their lives socially, i.e. through supporting independent living.

In other cases, the understanding of programme theory was exemplified in how participants described the way projects contribute to change through the specific *activities* that form part of the overall project or programme. One participant talked about how she had to use two different tools to engage with two different groups. When speaking about male group, she told me about tools for engagement *", the biggest thing being a pool table…. the ladies will sit around the table and chat really happily and for me I can find out what's going on and how I can support them. The guys are not into that whatsoever. They're not interested in sharing in that way but around a pool table, they will talk about absolutely anything. So, for me, that pool table is a massive tool of engagement"* [Org F]. In this example, the participant recognizes that, in trying to engage with the programme's target groups, she has to implement different activities for each group in order to engage them effectively. With respect to programme theory, this demonstrates an understanding that supporting change requires different activities for different groups.

Another example of the understanding of programme theory was in using understanding of change, at the level of the individual with whom the programme worked, to design services for others. One participant reflected on learning to work with individual young people through his employment programme. In talking about their approach to working with multiple individuals, he told me, *"it's just like whenever you're a child and you learn that you pull a handle a door opens, you sort of transfer that to other things. You can build up a profile, and because we're such a small group, we can see the same type of young person coming through and believe it or not, it's just like*

they're manufactured somewhere. It's like that's the exact same problem I've seen before, and this is probably what's causing it, and this is probably what will help" [Org E]. In this example, the participant is demonstrating that he can learn from what works with one individual and how he and his team can transfer that learning to other individuals. In terms of programme theory, this reflects an understanding of how the programme can help an individual and how using this understanding, at an individual level, can help to ensure change for other programme participants.

Whilst these findings show good implicit understanding of approaches to change through programme activities, this did not however appear to translate into more formal (systematic) evaluation activities. However, the participants from funding organisations stated that, whilst reporting mechanisms did not explicitly ask for a 'programme theory' of any sort, they had been structured around some key principles of programme theory e.g. activities (theory of action) and the difference made for the target population (theory of change). Participants from funding organisations also commented on the fact that there could perhaps be shift to funders taking the onus in gathering programme theory-type information. However, they were conscious of the additional pressures that would be placed on small TSOs, in particular, by requiring the use of programme theory. One participant from a funding organisation noted that, *"probably in the way that we ask applications to be written, or the information we ask for in evaluation, is kind of getting them to think through that* [programme theory]*, a bit more about what difference do you want to make, what will the activities you will do in order to achieve this be? It's there, even if not called a programme theory"* [Funder 1]

However, there was an acknowledgment amongst funders that asking for specific information from small TSOs, in the form of programme theory, can be burdensome for smaller organisations. Some participants from funding organisations reflected on approaches to gathering programme theory

information that placed the onus with the funder. On participant from a funding organisation told me that, *"we're piloting telephone calls with them to try to see what it was that was working … we want to do it with targeted programmes to unpick what was it that was most effective, what works and why and pull that out into learning points"* [Funder 2].

4.2.2.3 Using programme theory in 'unsystematic' evaluation activities

Despite the fact that a clear understanding of programme theory was not incorporated in what participants considered to be systematic evaluation activity, i.e. reports to funders, some participants reflected on their use of programme theory in unsystematic evaluation activity. In this sense, participants commented that their implicit understanding of how the programme or project was contributing to change was used within informal evaluation activities, i.e. used on a day-to-day basis to improve, change, and adapt services. One participant reflected on bringing staff together on a regular basis to talk about what it was that they were trying to change. She said, *"I try to encourage the staff to sort of say 'so what are we trying to achieve here?' Is it about communication skills or is it about bullying…and have we achieved anything? Have we modified behaviour?"* [Org A]. In this example, the project manager reflected on attempts within the organisation to get staff to talk about what it was being changed, i.e. behaviour, and through which mechanisms is was being changed, i.e. communication skills, all of which are examples of elements of programme theory.

Additionally, another participant similarly stated that he tried to encourage staff, through peer learning groups, to question what it was about the programme that ensured positive engagement with the young people. He told me that the reason was *"so that we can promote as much positive engagement with young people as possible, whether it be through activity or the relationship between the staff and young people"* [Org C]. Here, the participant reflected on the role of programme theory in ensuring that staff were constantly

engaging with young people in the most effective way possible. Overall, it appears that the use of programme theory, in unsystematic evaluation activity, is used to improve, change, and adapt services, and programme theory is not necessarily used in systematic evaluation activities.

4.2.2.4 Use of generic theoretical frameworks

Many of the participants commented on the use of 'generic theoretical frameworks' for guiding the development of programme principles as well as programme planning and design. Generic frameworks refer to guidelines from the specific field of work to which the programme relates, e.g. community development/youth work. Other sources of such frameworks are from government or local government. Generic frameworks tended to indicate how to ensure effective change within a given target group but on a more general level, rather than in the context of that specific programme, project, or service.

However, these cannot be considered as programme theory in the definition used in this study, albeit that such frameworks often include some aspect of 'theory': this is because the use of such generic frameworks do not relate to a specific programme in a specific context. For example, asset-based frameworks were mentioned by several participants. Asset-based frameworks state that positive reinforcement, alongside an asset rather than deficit focussed approach to working with communities and groups, should help to frame how organisations develop and deliver programme services (Hills, Carroll and Desjardins, 2010). Whilst the use of such frameworks was noted, some considered them to be of use at 'arms-length' only. Reflecting on the use of national and local authority community development guidelines, one TSO participant told me, *"I think as long as you use them as a guideline and don't get too involved in them because the thing about, for example, the locality, is that it's city-wide and there's a massive difference to what works here in AREA1 to say what perhaps works in AREA2. It's very much about using these things as a guideline but being able to identify how that links in with your*

particular area and not making it generic, but making it very tailored to the area that you work…I think you can get bogged down with a lot of the stuff sometimes but it is just about being able to pick out and identify the bits that work well within the area that you're in and can work to build that capacity" [Org F]. In this case, the participant recognised that such guidelines, whilst they have some value, only offer general guidelines which are not context specific for the programme. She acknowledges that in order to do effective work within a specific area, there needs to be an acknowledgement of local context, local challenges, and local needs.

4.2.2.5 Individual-level use of programme theory

One of the most common methods of formal, or systematic, evaluation mentioned was the use of case studies to illuminate the stories of individual programme participants. These case studies describe in detail the changes experienced by programme participants by being part of the programme or service, and were considered a good form of evidence by funders. Whilst, the understanding of change at the level of the individual, certainly contain elements of the how the theory of the programme is working for individuals, this level of detail was in most cases not considered at the level of the programme and in systematic evaluation activities of the programme as a whole. It was however used as evidence of the latter.

One participant reflected on how case studies capture the change in individuals by being part of the project. He noted specifically on how case studies are good to reflect, *"almost to an extent how that person feels, or where it came from, and what has changed … that has taken them to a better place mentally, or physically or whatever. I try and capture that"* [Org H]. In this example, the participant reflects on the use of case studies to capture what has changed, i.e. 'has taken them to a better place mentally, or physically', and how change occurs, i.e. 'where it came from'. Case studies reflect the use of understanding about how and why a programme works for an individual,

however it is not clear how this translates to consideration of how and why the programme is working more widely.

Another example of the use of individual level programme theory was in the integration of individual level programme theory within service delivery. In such cases, organisations used action plans to guide how the programme or project was going to help individuals. One participant commented on how they used individual training plans to ensure that each young person with whom they engage benefits from the programme. She said, "*We have individual training plans. 'This is what you're going to do, here's your action plan, this is how we're going to do it, when we're going to do it, why we're going to do it'"* [Org K]. In this example, the individual training plan in a sense, lays out an individual programme theory before the young person becomes involved with the programme to identify in what ways the young person will be involved in the programme and why, i.e. what are they going to achieve.

4.3 Overview of findings

This chapter presents the findings of the qualitative study, exploring the use of programme theory in small TSOs. In the first half of this chapter, I present the findings of this study relating to the context of programme evaluation practice in small TSOs. In particular, findings focus on the barriers and facilitators to doing evaluation, the 'systematic' and 'unsystematic' evaluation activities conducted, and the use of both the process of being involved in and noting the findings of programme evaluation activities.

With consideration of this programme evaluation practice context, the latter half of this chapter presents the findings relating to the use of programme theory in small TSOs. I find that whilst there is desire and need to use programme theory in systematic evaluation activities, both internal to TSOs and external for funding organisations, current systematic evaluation activities

do not make extensive use of programme theory. That being said, there are many instances, in the context of programme evaluation, that utilise the concepts underpinning programme theory. In particular, it was clear from talking with participants from small TSOs, that they have strong implicit understanding of programme theory and indications of its use in 'unsystematic' evaluation activities. There was also evidence of programme theory related ideas in the use of generic frameworks of practice and the use of 'individual-level' programme theory. Despite the evidence of its use in other aspects of evaluation activity, the use of programme theory did not seem to feature in participant reflections on systematic evaluation activities, despite this being of interest both internally within small TSOs and externally, to funders.

CHAPTER 5 STUDY 2 FINDINGS

5.1 Introduction

This chapter presents the findings of the second empirical study conducted as part of this thesis. The findings relate to the prescriptive research question addressed:

> *How* can the development of programme theory improve future programme evaluation practice in the context of small third sector organisations (particularly with respect to facilitating evaluative thinking)?

I present the findings relating to the learnings, emerging from the action research study, aiming to develop programme theory in two small TSOs, using the SODA methodology which was introduced in Chapter 3 (methodological framework).

5.1.1 Structure of findings

I divide this chapter into 3 parts. In the first two parts of this chapter I describe the development of programme theory using the SODA methodology in each of the two organisations respectively. This description includes: practical details on how I worked with each organisation; information on the programme being implemented in each organisation; details of the evaluation challenges faced within each organisation; the approach to developing programme theory using the SODA methodology; and key learnings emerging from each application. The last section of this chapter compares the action research approach, and the emergent learning in each organisation, before summarising the mutual learning points across both organisations.

It is worth noting here that I was careful to ensure that the learning presented, with respect to each organisation, was closely related to the development of the programme theory. I did not wish to focus solely on the added value, to participating stakeholders, of having dedicated time to talk about programme

evaluation challenges. It follows that the emerging learning presented relates primarily to the process of development and content of the maps, and the subsequent discussions about the specific evaluation challenges facing each organisation.

5.2 Developing programme theory

5.2.1 Organisation 1

5.2.1.1 Working with Organisation 1

I first came into contact with Organisation 1 whilst completing my MRes degree and during the first months of my PhD. At that point I was exploring the economic consequences of youth disadvantage and effective strategies to target youth disadvantage. Organisation 1 is a third sector organisation based in Scotland that has developed, and currently implements, a school-based mentoring programme to target some of the impacts of youth disadvantage. The school-based mentoring model is an internationally used strategy to alleviate some of the consequences of youth disadvantage. For example, the Big Brothers, Big Sisters[21] mentoring programme has been widely applied in the United States and in Ireland.

Organisation 1 was in the process of developing more formal research partnerships within the university; however, I took the opportunity, with the guidance of my supervisor, to communicate to the programme management team, my research and methodological interests, based on the evaluation challenges they identified. Out of this grew the proposal for this piece of action research. At the very beginning of the proposal development period, I was focussed on programme theory from the perspective of generating learning about how and why the programme model worked. However, as the months went by, my research focus and interests developed to focus more on the practice of evaluation within third sector organisations, and how to support this

[21] http://www.bbbsi.org

practice. In the end, I proposed this piece of action research at the end of the first year of my PhD: we then proceeded with the development of programme theory, which I outline in more detail in the following sections.

In general, working with the organisation involved communicating via e-mail with programme staff (primarily the management team). However, I also had the opportunity to be involved in other aspects of the programme including: attending a conference held by the organisation; conducting site visits to schools to hear about the programme from full-time members of staff; meetings with the programme management team to hear about programme developments and challenges; and participating, alongside programme management, in a trans-Atlantic seminar with other researchers who conduct research on mentoring programmes, and other aspects of the mentoring relationship. This rich involvement with the programme provided me with good contextual understanding of the programme whilst conducting this action research study.

5.2.1.2 The programme

With a base (at the time of conducting this study) in over 30 secondary schools in the Glasgow area, the core component of the programme is school-based mentoring. Volunteer adult mentors spend 1 hour per week with a young person with whom they have been matched. Young people are referred to the programme on the basis of several criteria relating to their experience of the care system and their perceived vulnerability. At the point of referral, young people can then choose to opt-in to participate in the mentoring component of the programme or decide that they do not want to take part.

The core idea underpinning the programme is that, through the relationships mentors built with young people, they can offer them support and guidance for the challenges they face. The overall aim of doing so is to close the attainment

gap amongst advantaged and disadvantaged young people in Scotland by enabling them to realise their potential. Other components of the programme aim to further support this goal and include: group work to familiarise younger pupils with the programme before participating; half-day work experience opportunities to give young people an introduction to various career paths and job opportunities; and an ambassador programme for a small number of young people to represent and develop the programme on behalf of their peers. A full-time member of programme staff is based within each participating school to co-ordinate the programme at that site.

The overall concept of school-based mentoring is a well-researched strategy to working with disadvantaged young people in schools (Wood and Mayo-Wilson, 2011). The ideas and concepts behind developing relationships between volunteer mentors and young people and the support they can offer are, as a result, well-developed concepts in the wider academic field of mentoring. Even so, well-defined strategies to targeting social issues, are context dependent and as such, require careful contextual considerations in implementing such strategies. The programme developed and implemented by Organisation1 had taken particular account of contextual issues relating to youth disadvantage and attainment, with many programme staff members having worked in the education/youth disadvantage sector prior to becoming involved in this particular programme.

5.2.1.3 Identifying evaluation challenges

In several meetings with programme management, I, alongside other academics from both the business and humanities and social science faculties, were able to discuss with programme management some of the evaluation challenges being faced. It emerged that because the organisation was going through a period of rapid expansion, upscaling the programme to other locations across the country, there was interest in understanding more about how and why the programme was working for the young people i.e. better

understanding of the programme theory. They anticipated that a better understanding of how and why the programme was working for young people would enable them to understand how to adapt, improve, and apply the implementation principles of the programme on a wider scale. Moreover, it was hoped that this understanding of programme theory would also support more appropriate and effective programme evaluation through greater understanding of young peoples' perspectives.

The organisation already had a logic model (programme theory) that reflected how the programme was assumed to work, based on both existing school-based mentoring theory. The logic model was also informed by the contextual impact of the challenges and barriers facing disadvantaged young people attempting to realise their potential in Scotland. However, the logic model was developed in the early phases of programme development and implementation, and had not been updated since. They captured information relating to programme theory in the form of individual case studies, but they were interested in understanding, more systematically, what it was about the programme that helped the young people. They found that it was challenging to generate systematic learning from across individual case studies. Therefore, they became interested in systematically capturing the perspective of young people.

5.2.1.4 Approach to developing programme theory

Deciding on the approach

It had been decided with programme management that there was a requirement for a better understanding of the how and why of programme functioning, from the perspective of participating young people. As such, I proposed the SODA mapping methodology to programme management. After discussing the SODA methodology, and what it could and could not help the organisation achieve, we decided that using individual young people's perspectives, to 'map' experiences of the programme through a cognitive

mapping approach, was the way to proceed. Together, we devised an approach that entailed conducting individual interviews with young people, mapping those individual perspectives of the programme, and using the analysis of the individual maps to conduct facilitated discussion with programme management: the purpose of a facilitated discussion was to discuss the key learning points emerging from the mapping process as well as what these key learning points meant as they expanded the programme. We also discussed using a focus group format, however given practical time constraints, it was decided that this approach was less feasible.

I conducted some short pilot interviews with six young people to understand the kinds of questions young people were comfortable answering, and to address how the mapping process might work in real-time. The pilot interviews were conducted in the summer of 2017, during an ambassador training workshop: this workshop trained young people to participate in the ambassador component of the programme. I developed a question guide, adapted from a semi-structured interview protocol for developing logic models (Gugiu and Rodríguez-Campos, 2007). Questions in the pilot question guide focused on their experiences of the programme, what young people felt that they achieved by participating in the programme, and factors affecting their participation. I used the pilot interviews to trial whether it would be feasible to conduct the mapping in real-time with the young people.

In general, young people in the pilot interviews responded well to the questions with the only additions being the inclusion of an opening question to make the interview process more accessible. In addition, staff were consulted regarding wording of questions and appropriateness of the protocol. In conducting the pilot interviews, I realised that it was better to give my full attention to the young people in the interview rather than be focussed on mapping in real-time. The reasons were to ensure that young people were engaged and not intimidated by the interview process and because I felt I was not yet experienced enough

in SODA mapping to do it effectively in an interview setting. Based on the pilot interviews and consultations with staff, I developed a question guide (Appendix 4) and decided that the mapping would be undertaken post-interview to enable me to focus on the conversations with young people. We decided the best approach would be to first, conduct the mapping and do some preliminary analysis, and then to bring key learning points to the programme management team in a facilitated discussion session.

Developing the programme theory

The defined approach entailed interviewing 18 young people who were participants in the programme; these interviews were carried out between March 2018 and July 2018. For practical reasons, these young people were participating in the ambassador component of the programme which meant interviews could easily be conducted outside of school hours during the ambassador training sessions. Ambassador training sessions aim to bring ambassadors together in a school setting, but outside of school hours, to discuss the challenges facing young people and, based on this, develop and improve the programme. Holding the interviews during the ambassador training sessions meant little disruption to the school day of the young people. Young people were first asked if they would like to participate by a member of programme staff responsible for the ambassador training component of the programme. At this point, participants were given a participant information sheet (Appendix 5) and consent form. Participants under the age of 16 also required parent/carer consent to participate.

The semi-structured interview format allowed me to address key questions of interest regarding the programme, its outcomes, and factors influencing programme outcomes, whilst affording the flexibility to probe areas of interest mentioned by the young people. The semi-structured interview protocol focussed specifically on young peoples' experiences of the programme including their relationship with their mentors, factors influencing their

experience of the programme, and their perspectives on what changes they felt they had experienced as a result of being involved in the programme. The question guide used is attached in Appendix 4.

In general, interviews lasted between 30 minutes and 60 minutes and were audio recorded and transcribed verbatim. Maps were built by the researcher based on both the transcripts and the audio recordings. Following the mapping conventions of the SODA methodology, it can be beneficial to map in real-time during participant interviews, however, based on my experiences in the pilot interviews, I felt it was important for me to be able to offer my full attention to the young people to ensure they felt comfortable talking with me: in this sense, I was wary of being distracted by mapping as I was not yet so experienced in its use. I did however use small word and arrow diagrams, following the means-end mapping format during interviews to help me clarify some meanings with young people. Moreover, in cases where young people were struggling to answer some questions, I used word prompt cards which showed words relating to the programme with which participating young people would have been familiar with. In cases where, for example, a young person was struggling to answer a question, I asked the young person to pick some word prompt cards to help them try to articulate what they wanted to say. Some examples of words used in word prompt cards are shown in the word cloud in Figure 7. The prompt cards where only used in 2 interviews.

Figure 7 Examples of words used on prompt cards for Organisation 1 interviews

As interviews took place within schools, I also had the opportunity to meet and speak more informally with programme staff based in schools, programme management, as well as informally with other participating young people and their mentors. This informal contact enabled me to do two things. First, I could report in real-time to other programme stakeholders as to what the aim of doing this research, and the implications it could have for the programme. Second, more informal communication with programme stakeholders gave me a better idea of programme operations on a day-to-day basis and the challenges faced by those delivering the programme. By being more aware of the day-to-day implementation of the programme and the perspectives of other stakeholders, I was able to be more sensitive to the challenges faced by both programme management, and those 'on the ground', and incorporate this understanding into my approach to working with this organisation.

Map development, analysis, and validity

Upon the completion of interviews, audio recordings were transcribed verbatim and maps were created in the *Decision Explorer* software. Maps were created

based on a mixture of audio recordings and transcriptions. However, for reference, I have briefly described some of the steps taken here. First, the use of the audio enabled me to understand better, and contextualise the meaning of specific statements with respect to the tone of voice or the way in which the statements were made. However, the transcriptions provided a good point of reference to cross-check all relevant details were included in the maps. Once maps were created, I conducted a series of analyses on individual maps to identify interesting features and themes. I then used this to identify whether these features were common across the maps. For example, I used the 'concept style function' on Decision Explorer to differentiate between different programme components, outcomes, and negative influences (these are illustrated in Figure 8). The concepts in individual maps were coded using these concept style functions which meant that the map could be easily visually inspected through looking at different uses of each concept style. Other analyses on individual maps included which proved useful were identifying 'heads' or outcomes in each map, and identifying 'busy' or central concepts.

Figure 8 Organisation 1 - concept style examples

The features highlighted in the analysis, considered alongside that which I had learned informally about the programme, were then used as the basis of the facilitated discussion. An important point to make here is that the individual maps did not go through a process of validation with each individual young person, i.e. I did not return to each interviewee to check whether the map I had produced accurately reflected what they said. This was a conscious decision for a number of reasons. First, the time required to go back to each individual young person could be considered disruptive to their school day/ambassador training, given that they had already given up some of their training time to speak with me. Second, as this mapping exercise was not necessarily being considered or used for any formal evaluation activities, the team did not consider it necessary to go back to the young people for validation. Had this been a formal evaluation activity, I argue that it would have been necessary for me to validate each map with the young people. Third, in the interviews, I was conscious to clarify anything which was unclear to me at the time in order to ensure that maps accurately reflected the meaning which young people attached to concepts, even if only in retrospect.

Facilitated discussion

An important part of the action research approach was to host a facilitated discussion towards the end of the study. The purpose of the facilitated discussion was to discuss the key learning points emerging from the mapping, and to think about how these key learning points could be taken forward as the programme continued to expand, and in terms of programme evaluation.

The facilitated discussion took place in the head office of the organisation in March 2019 and involved 6 key programme stakeholders (+ 1 unable to attend on the day). I and another PhD student were also in attendance. The discussion involved a brief explanation of the mapping process and the analysis of maps to remind and re-familiarise staff with how the interviews and subsequent mapping were carried out, and so as to maintain process transparency. This explanation was followed by a participatory discussion of key learning points emerging from the mapping process and then a longer discussion on what this meant in terms of the evaluation challenges they faced, i.e. the expansion of the programme. One key member of the management team was unable to attend on the day of the facilitated discussion, so I carried out a similar, but shortened, session with her the following day. This ensured the participation of the whole programme management team who were involved from the beginning.

Discussing the maps and map content

After providing an overview of the approach I had taken to create the maps, I summarised three key learning points from the content of the maps using the visual maps to facilitate and support discussion around these points. The learning points discussed were jointly based on the map analyses, as well as the saliency of learning points linked to the initial evaluation challenges that had been discussed with programme management, alongside my

conversations with programme staff, beneficiaries, and management throughout my involvement with the programme. The three key learning points I highlighted in the facilitated discussion centred on: improving career-based outcomes for young people; the perceptions of positive mentor matches; and influences on young peoples' decisions to participate in the programme. I discuss each of the key learning points in turn below, using excerpts from maps to illustrate what the maps looked like.

1. Most of the young people reflected on how their involvement in the programme helped them to follow, or think about, their desired career path. An important goal of the programme is to close the attainment gap for young people and so, I considered this to be a salient feature to explore with programme management. In many of the individual maps 'follow my chosen career path after school' (or similar) can be found as a head in the map. I used the 'heads' function on Decision Explorer to list all of the heads (or outcomes) in each individual map. The head, or outcome, appearing most frequently at the top of individual maps was with respect to being able to follow or pursue a desired career path after school. I also mapped all the heads across the individual maps, in a meta-outcome map, to get a sense of how the outcomes of the programme were structured according to young peoples' perspectives (as this map is too big to include in the main body of this thesis, I have included it in Appendix 6. Whilst the individual concepts cannot be read, Appendix 6 aims to give the reader a sense of the complexity and richness of information collected regarding outcomes).

   From across each of the individual maps, we were able to look at all the different ways in which young people perceived their involvement in the programme helped them to do follow their chosen career path. To explore this idea in more detail, I used the meta-map for heads (Appendix 6) and simplified it such that a second map showed only those concepts relating to pursuing a desired career path. This map is

shown in Figure 9. This map could then be used to discuss the outcome for all participating young people in exploring the various ways in which they perceived the programme was helping them to reach the goal of pursuing their chosen career path. According to the map, some of the key ways through which young people perceived the programme supported them follow their desired career paths relate to: increasing their desire to work in a given area; belief in themselves that they can do and be what they want; belief in themselves that they can do well; and being prepared for following their desired career path after school.

Another interesting feature of this map for programme management was the use of colour coding to indicate visually how the various programme components were perceived to help young people achieve goals (*facilitated discussion notes, 4th March 2019*). For example, blue here refers to some aspect of the mentoring component and brown refers to the work experience component. This allowed programme management to see the how the young people recognised their mentors as playing central and influential roles in assisting them to determine and pursue key, chosen career paths.

2. From informal conversations with programme staff in schools, it became clear that the early stages of the mentoring relationship were particularly volatile. Therefore, understanding the factors that aid positive relationship formation, at an early stage of the mentoring relationship, was a salient feature of the interviews to explore in map form. Programme staff felt that improving their understanding of this would enable more effective 'mentor-matching' from an early stage. I had asked young people about their relationships with their mentors and what made them decide at the beginning of relationship that this was a person they would like to spend time with. As such, I pulled this information from the individual maps to create a 'merged map' on positive mentor match perception. I mapped all the different factors

young people perceived to be associated with a positive mentor match in the early stages of the mentoring relationships. Figure 10 shows the map reflecting the various perspectives on what makes a positive mentor match. The numbers at the top of the map, attached to dotted lines, are the consequences of a positive mentor match. For the purposes of discussing what makes a positive mentor match, these concepts were 'hidden' from the main map.

What was interesting for programme management, from this map, were the wide variety of factors that influence early perceptions of a positive mentor match: this is indicated by the width of the map. In particular, the confusing role played by the mentor's age (shown in the concepts boxed in red in Figure 10) was useful information for programme management to consider in the matching process. According to the young people I spoke to, the mentor's age had both multiple positive and negative impacts on positive match perception. Whilst there was no obvious solution to this challenge, programme management considered it valuable information to put into context why some young people might prefer an older or younger mentor *(facilitated discussion notes, 4th March 2019).*

3. One of the features emerging from the individual interviews, and from my conversations with programme staff, was the decision to participate in the programme (given that young people are referred but need to 'opt in', i.e. make the choice themselves to participate, post-referral). It follows that all the young people I interviewed had decided to participate. So, instead of mapping this in individual maps, I decided to map the decision to participate as a separate map from the beginning. The analysis of 'busy' concepts[22] in this map highlighted two particularly

[22] Those concepts or constructs with the most arrows pointing in to them, i.e. those concepts with many factors influencing them.

important factors on the decision of young people to participate These are: (1) the belief that the mentor can change the young person's life, and (2) being approached to participate in the programme, rather than going unnoticed by the programme. As the larger map on 'decision to participate' is too big to include in the main body of this thesis, I have included it in Appendix 7. The two 'busy' concepts, boxed in red in Appendix 7, are shown separately in Figures 11 and 12. Looking more closely at these two factors, we can identify several factors which are important to consider. For example, in order to achieve the 'belief this person can change my life' (Figure 11), young people perceived that it was important to clearly understand the role of a mentor as well as understanding the potential benefits of the programme. Moreover, there were some negative influences on the 'belief that this person can change my life' such as 'not wanting a mentor' and 'the belief that you will be ok without a mentor'. Overall, this analysis facilitated a closer look at all the factors influencing young peoples' belief that a mentor can change lives. The same can be said for Figure 12.

Overall, this analysis was of interest to programme management because they acknowledge that there are groups of young people who the programme cannot reach because those young people decide that they do not want to participate in the programme. One of the challenges they faced was encouraging those who did not opt-in to do so. Using information, such as that discussed above, can better inform strategies to encourage participation, particularly in identifying where there might be doubts for young people concerning their participation in the programme, and about the kind of information about the programme that is valued by young people.

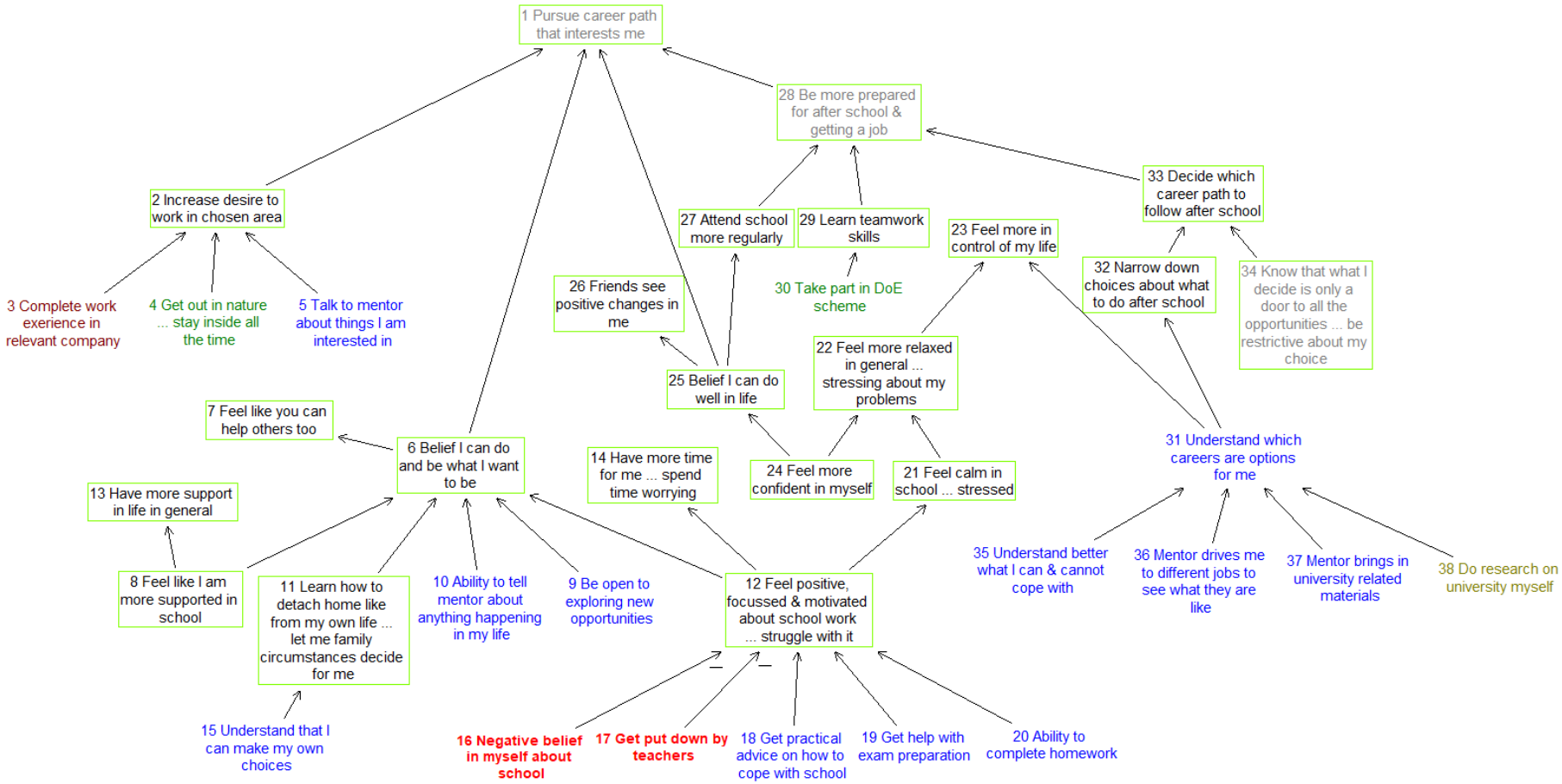


Figure 9 Simplified map of 'desire to pursue chosen career' outcome

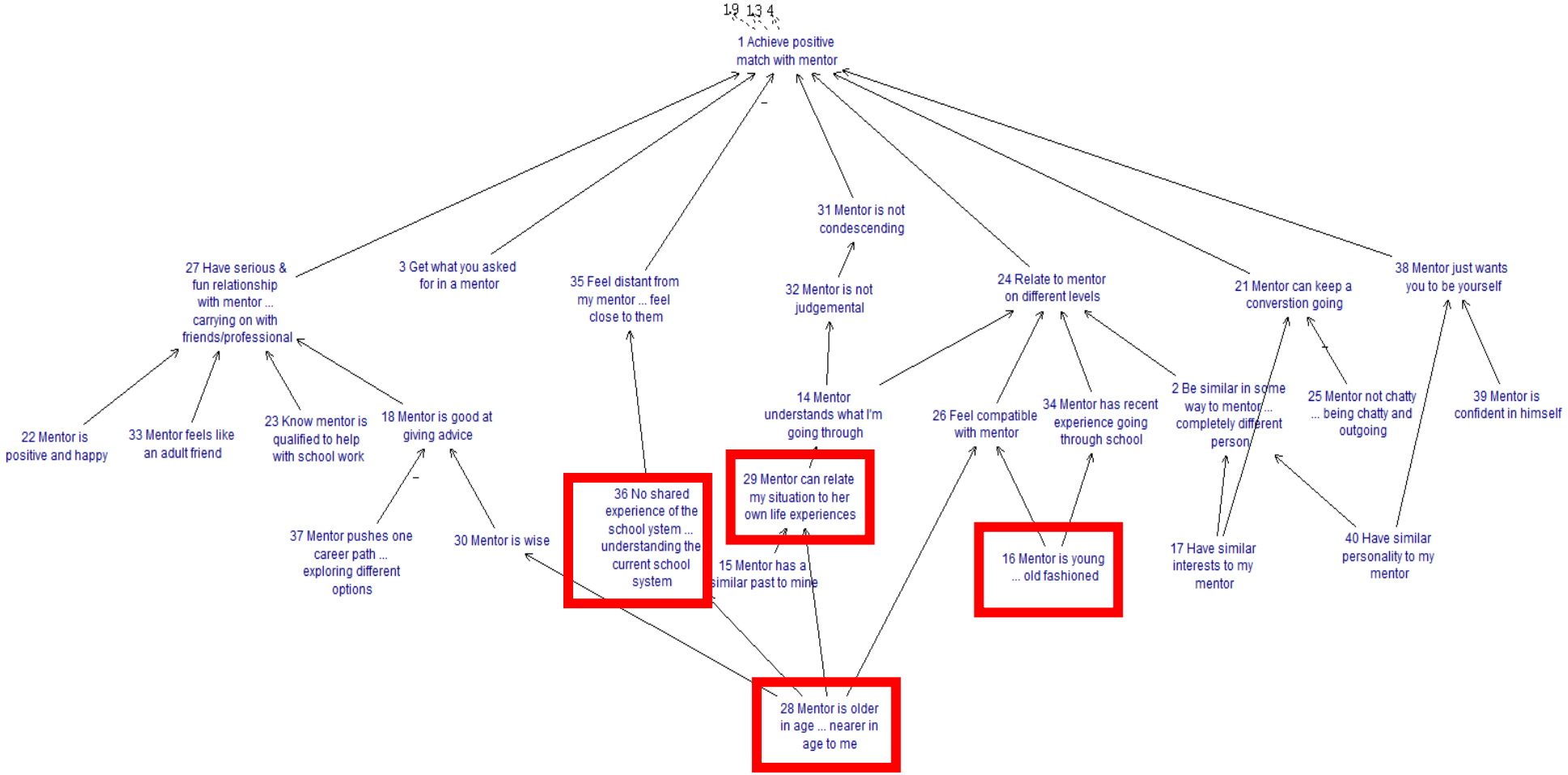


Figure 10 Map of means of perceiving an early positive mentor match

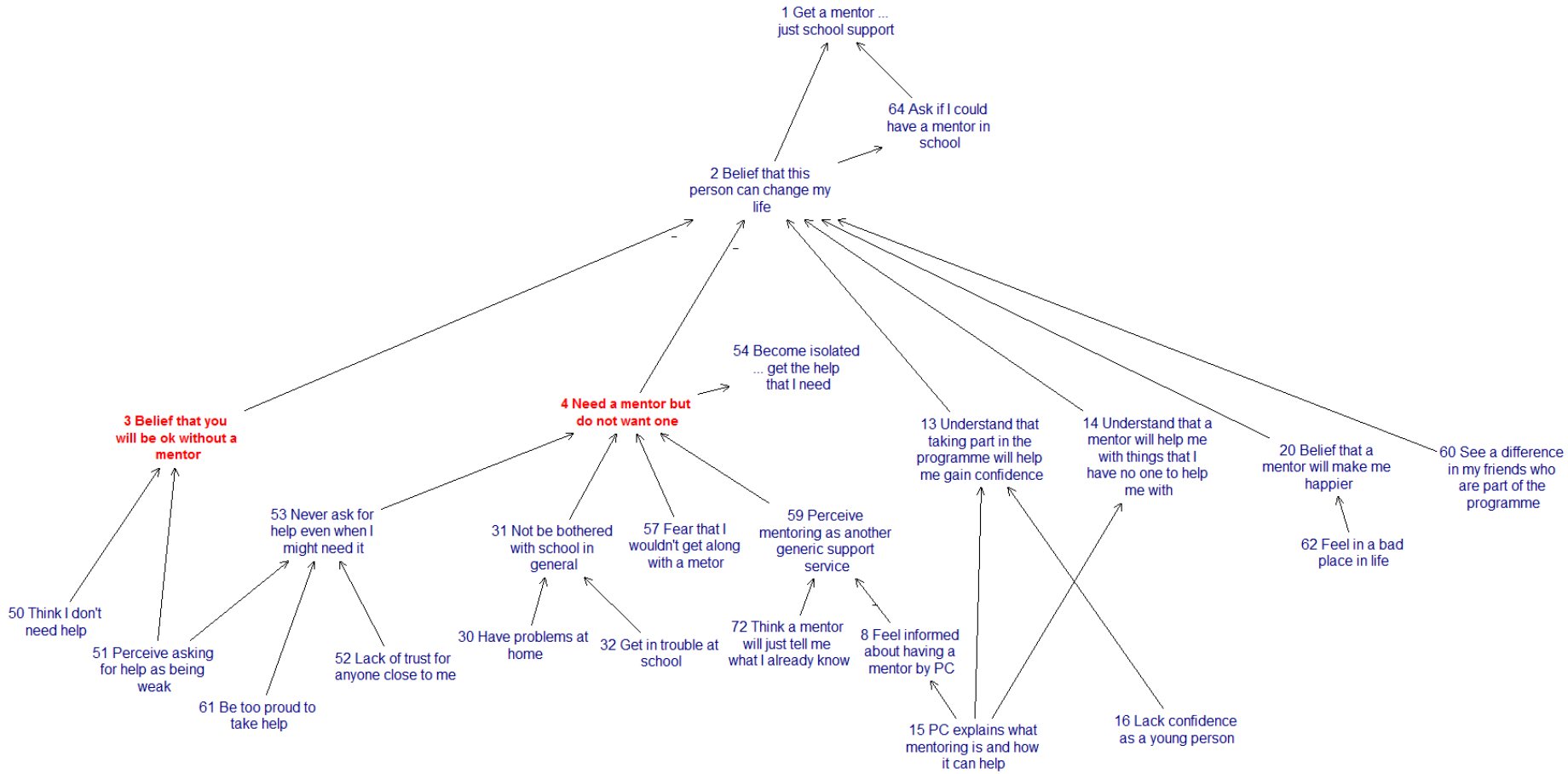


Figure 11 Zoom 1: Influence on decision to participate - belief that this person can change my life

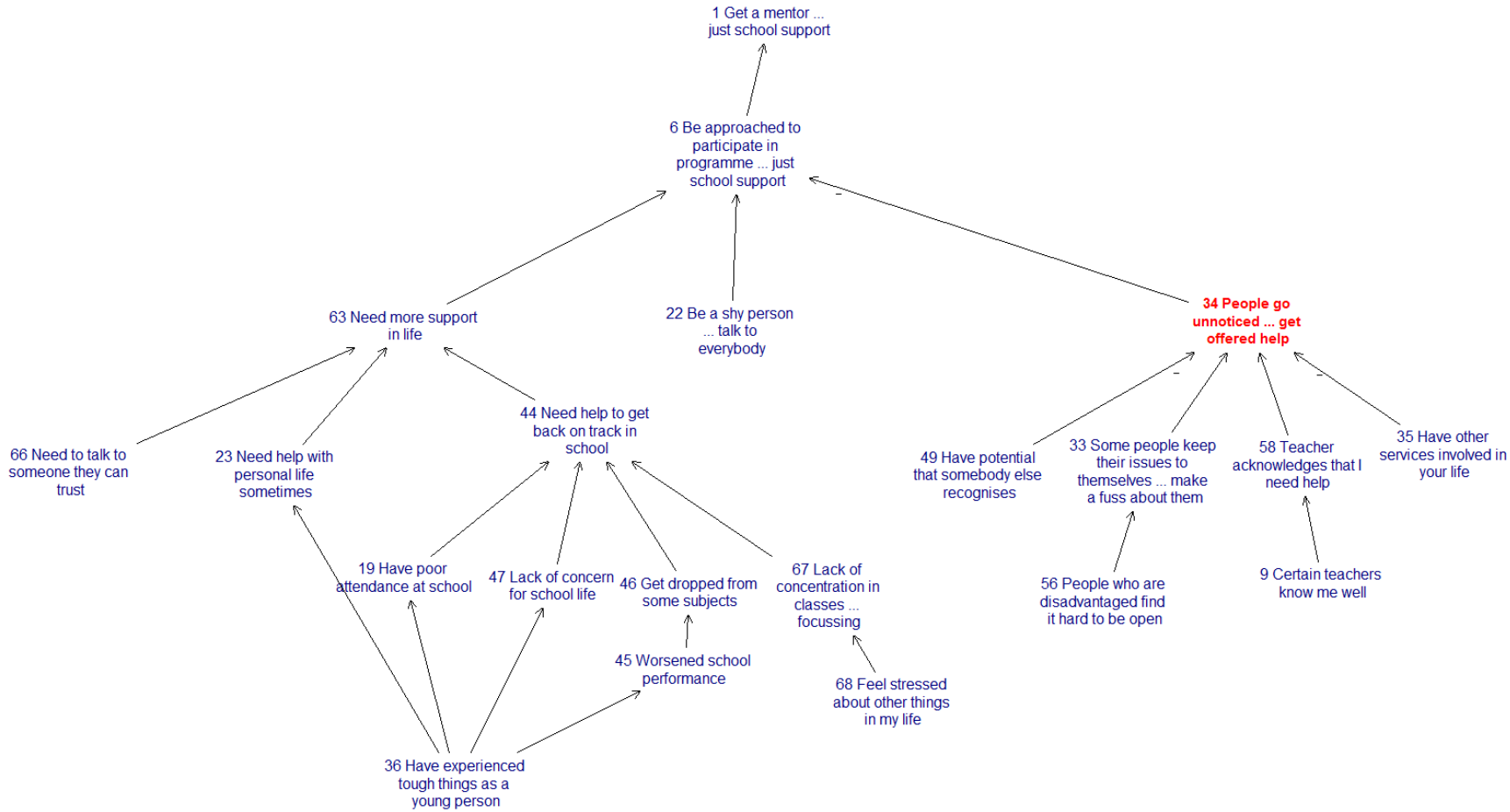


Figure 12 Zoom 2: Influence on decision to participate - be approached to participate in the programme

5.2.1.5 Key learning

The discussion of the three key learning points described above facilitated discussion about their significance in terms of the evaluation challenges faced by Organisation 1: these challenges included the rapid expansion of the programme and the need to understand programme theory from the perspective of participating young people so as to support future programme implementation and evaluation. Learning emerging from developing programme theory in Organisation 1 is presented with respect to the learning emerging from facilitated discussion with the maps, the implications for programme implementation, and the implications for conducting programme evaluation.

Learning from facilitated discussion with maps

In the case of developing programme theory in Organisation 1, participating programme management felt there was value in the process of facilitated discussion in terms of interacting with the maps, because of map content, and the visual nature of the maps. It was this interaction that facilitated the learning that I present in the following sections on 'implications for programme implementation' and 'implications for conducting programme evaluation'. However, I am cautious about pinpointing exactly what it was in this approach that had value to the participating stakeholders in terms of programme implementation and conducting programme evaluation. For example, I could question whether it was the visual nature of the map or the content that stimulated the learning I present in the following sections. The learning I present thus reflects what I have interpreted to be most closely related to the development of the programme theory, rather than simply having dedicated time to think about programme evaluation.

Overall, with respect to the content of the maps, participants concluded that the content of the maps was not particularly surprising. Even though there was considerably more detail in the maps than they had been able to capture

before, the key learning points emerging from the content of the map were in general unsurprising to the programme management team. One could argue therefore that the content emerging from the mapping process was not so useful as it was simply highlighting what the programme management team already knew.

However, members of the programme management team noted that what was of value, was the discussion stimulated by having the familiar information presented in this way. Having information presented in a *visual way* enabled the team to see familiar information in a different light. This stimulated questions such as: “we already knew that, why haven’t we done anything about it?” *(facilitated discussion notes, 4th March 2019)*. The programme team commented on the fact that the visual nature and structure, particularly colour coding and looking at ‘busy’ concepts, facilitated their ability to think differently about things that they already knew. The maps also stimulated careful analyses by offering more systematic information about young peoples’ experiences at hand. One member of programme management commented on this, stating that, *“for a visual learner like me, it helped me make sense of things better than individual case studies” (Programme Development Manager, feedback from facilitated discussion, 4th March 2019).* In this sense, it seems that the visual, tangible nature of the maps provoked thinking and discussion beyond that which the approaches that they currently use to address these issues e.g. case studies can yield.

Moreover, even though maps presented information which was unsurprising to the team, participating members of the programme team also highlighted this favourably as it confirmed their understanding of how the programme was working with more systematic detail on why this was so. For example, as I described in the previous sections, an acknowledged challenge for this organisation was encouraging those who, in spite of being referred, chose not to participate in the programme. Whilst they already knew this was a challenge,

little action had been taken to address the issues with respect to the 'opt-out' rate, of some young people, for a number of reasons. But programme management commented on how having more systematic information on these issues presented in a visual way, helped them to think a little differently about these challenges and the necessary strategies to address them *(facilitated discussion notes, 4th March 2019).*

Implications for programme implementation

The key learning points emerging from the visual maps described above facilitated discussion on how the programme was being implemented and ways to improve this. Participating members of programme management were able to think more purposefully about how implementation of the programme could be improved, compared with how it was being implemented currently, or how they thought it was being implemented. A specific example of this was being able to understand in more detail the reasons why young people made the decision to participate in the programme (as discussed in the preceding section on *'discussing map content'*). This stimulated discussion on potential strategies to encourage those who do not 'opt-in' to the programme participation. In this case, some members of the management team noted that the use of the young people's language in the map and findings was particularly illuminating *(facilitated discussion notes, 4th March 2019)*. They thought that the use of young people's language made it easier to understand young people's perspective on participation and encouraged them to think differently about the challenge of involving those who did not see the programme as useful.

Another example of implications for programme implementation was in better understanding the early perceptions of a positive mentor match. Programme management felt that this kind of information could be used to inform the future development of mentor training materials. In this respect, training materials could be developed to better account for what young people perceive, at least

in the short term, to be important factors of perceiving a positive mentor match e.g. mentor age. In both of these examples, programme management were again able to think more critically about current programme implementation and to develop thinking around potential strategies for strengthening and improving implementation *(facilitated discussion notes, 4th March 2019).*

Implications for conducting programme evaluation

The key learning points discussed above also stimulated discussion about future programme evaluation, especially with respect to data collection approaches and resources within the organisation. For example, whilst there was enthusiasm to extend the mapping study and a desire to learn more about the programme from a wider group of young people, it was recognised that this kind of work was resource intensive and would require additional resources which were not readily available to the organisation. In light of understanding research resource constraints, programme management also discussed in detail how the key learning points from the mapping exercise could better support and develop current data collection, monitoring and programme evaluation processes, rather than conducting additional research studies. For example, the group discussed how they could use more detailed qualitative understanding of the programme to support more targeted quantitative data gathering. Specifically, the group talked about how to better collect data on mentor relationship characteristics and explore how this might be linked to experiences of the programme *(facilitated discussion notes, 4th March 2019)*.

Overall, the discussion with respect to programme evaluation centred around how, given resource constraints, the learning points emerging from the content of the maps could support the development of more effective and informative data collection processes that fit within the existing data processes. The discussion about how to support the development of better data collection processes points to an acknowledgement on behalf of programme management, of a need to investigate how to improve their capacity for doing

programme evaluation, and to think more realistically about their information needs.

Summary of learning from Organisation 1

I developed programme theory in Organisation 1 due to the identified need to better understand how and why the programme they delivered was working, from the perspective of participating young people. The need for understanding more about how and why the programme was working arose as the programme was undergoing a period of rapid expansion and those delivering were conscious of forthcoming implementation challenges, as well as their ability to evaluate the programme.

Using the cognitive mapping approach and the SODA methodology, I mapped young people's experiences of the programme and used these maps, and their analyses, to conduct facilitated discussion with programme management. Whilst the participating stakeholders found that map content was in general unsurprising, they noted that there was value in systematically capturing in more detail on the programme, offering multiple perspectives and the potential for visual learning. Alongside the increased detail, the visual nature of the maps and the use of young peoples' language, illuminated some challenges that had previously been acknowledged but not been addressed. Furthermore, the opportunity to partake in the facilitated discussion on the maps provided the space and means for participating programme management to think both critically and purposefully about programme implementation and outcome evaluation. In particular, participating programme management were able to think about how current processes, both in terms of programme implementation and outcome evaluation, could be improved, in light of their existing systems without placing too much additional burden on the organisation.

5.2.2 Organisation 2

5.2.2.1 Working with Organisation 2

I first came into contact with Organisation 2 during the 2nd year of my PhD. Organisation 2 are a third sector organisation based in Scotland, with headquarters in Glasgow. As part of the community-based arm of the organisation, they had developed and were implementing a programme for young people that simultaneously trained them in first aid, peer leadership, and facilitation skills. A board member of the organisation had been put in touch with some of the staff in my department regarding conducting Social Return on Investment Analysis (SROI).

Organisation 2 was seeking some guidance from the university in conducting SROI analysis to identify the social value of their programme. They wished to use this to generate more funding to upscale the programme. During an initial meeting in February 2018 with 4 programme stakeholders, my supervisor, another PhD student, and myself, we discussed SROI alongside some other evaluation related challenges they were facing. We discussed the complexity and resource intensive nature of SROI, as well as Organisation 2's current programme evaluation processes. The programme stakeholders told us that, for this particular programme, their evaluation processes were in the early stages of development, as the programme had only been implemented in one site so far. At this point, they had no real working framework for conducting evaluation or an explicit theory of how the programme worked. So, from here, with the guidance of my supervisor, I communicated to the team that this was my area of study and that it might be worth thinking about developing a framework, such as programme theory, before thinking about conducting SROI analysis. This kind of framework could then serve as the basis for future efforts to do SROI given that developing programme theory is a key part of conducting SROI. Based on this suggestion, I proposed that we meet as a group in the coming months to discuss developing a programme theory in more detail.

In the intervening weeks, and months, I was able to learn a bit more about the programme via e-mail communications with the programme stakeholders and some documentation they had on the programme and its evaluation, including the organisation webpages. I held a follow-up meeting was held in March 2018 including the programme team and another PhD student. In this meeting we were able to spend some more time discussing the concerns of programme stakeholders regarding programme evaluation. At this point, I was able to introduce my research interests in more detail and explain how there were some synergies between the challenges they were facing and much of what I was interested in. In this meeting, we collectively agreed that exploring the idea of developing programme theory might be a useful place for them to start, with respect to evaluation and building a basic evaluation framework. From this point I was able to develop the proposal for this application of the SODA methodology. I detail the specific approach in the following sections. However, I communicated this workshop proposal to programme stakeholders, all of whom agreed with the details. From here, we arranged a time and date for the workshop and worked collaboratively to generate a workshop schedule. The workshop was held in April 2018.

5.2.2.2 The programme

Organisation 2 has developed and implemented a programme which involves the delivery of first aid training alongside a peer leadership and facilitation course to small numbers of pupils in Scottish schools. In the early stages, this programme delivered only first aid training to over 600 pupils. However, the programme now implements an additional component. The addition of a peer leadership and facilitation element enables a smaller number of trained young people to then deliver first aid information sessions to other pupils within their own schools. The programme also includes training for teachers to support young people to use their peer leadership and facilitation skills once they have finished the programme. The first aid qualification gained is certified for 3

years. The peer leadership component of the programme aims to provide the smaller number of participating young people with wider employability skills. Through training young people in the skills needed to deliver first aid training to their peers in school, the programme aims to improve self-confidence, leadership, and presentation skills. At the time of completing the action research, the programme (including the peer leadership and facilitation component) had trained just over 30 pupils.

5.2.2.3 Identifying evaluation challenges

In an initial meeting with key programme stakeholders, we discussed some of the challenges facing the organisation in terms of the programme and its evaluation. The organisation sought to increase the reach of the programme through increasing the number of schools in which the programme was being delivered and also through developing partnerships outside the education field, including partnerships with organisations supporting young carers and those in the criminal justice system. The programme team saw the potential value the programme model might have with these groups of people and in these settings.

In order to do so, they had to secure funding and new partnerships, for which they would be required to conduct an evaluation of the current programme in order to show it to have benefits for the targeted young people. The organisation was interested, in this respect, in better understanding the social impact of the programme. Employees of the organisation also had questions around the Social Return on Investment, and the alignment of programme outcomes with Scottish Government policy objectives in relevant areas such as community development and safety. The programme stakeholders considered the programme to be innovative and, as a result, it was difficult to locate examples of evaluations of such programmes to help them. This left the programme team a little stuck with how to approach programme evaluation. At the time of the meeting, the organisation therefore had no plans for how they

would evaluate the programme, other than an early draft of an evaluation plan which included some key questions around the needs of the target population, programme planning, programme implementation and programme outcomes *(e-mail communication, 9th March 2019)*. They acknowledged the need to think about these plans more formally to create a framework that would allow them to gather the information they needed in order to secure the funding for programme expansion and thus, the programme's long-term sustainability.

5.2.2.4 Approach to developing programme theory

Deciding on the approach

In a second meeting with the programme team, the lack of evaluation plan served as a basis for discussion. We discussed what developing a plan might mean in the context of this programme. I suggested that developing a programme theory might be a useful starting point to enable the organisation to consider what they wanted to evaluate and how they would go about evaluating it. This would involve thinking about the links between the programme that they deliver and the outcomes they sought to change.

From this point, I spent some time drafting a proposal for an Oval Mapping workshop. The reason for proposing an Oval Mapping workshop was the fact that the organisation had never developed the programme theory explicitly before. I proposed that the 'quick and dirty' format of the Oval Mapping workshop would be useful to surface some initial thoughts and assumptions on the programme theory in a relatively short space of time. From there, it would be easier to consider more detailed work and analyses if need be. Upon communication of this proposal to the programme team, we agreed that this was how we would proceed. A date and appropriate working space at the organisation's office were arranged by the programme team. The programme team also decided who, up to a maximum of 10 people, would participate in the workshop. I left this decision in the hands of the programme team as I felt

they would be better placed to decide on who could be best to contribute to the workshop in a meaningful way.

Developing the programme theory

Developing the programme theory in this setting involved the use of the Oval Mapping Technique in a facilitated half-day workshop with 6 programme stakeholders including:

1. A grants manager with responsibility for writing and submitting applications for programme funding,
2. A programme manager who has the responsibility for both developing and overseeing the delivery of the programme,
3. 2 Training managers who have the responsibility for developing training workshops in the organisation
4. A trustee who has the responsibility for sitting on the board of the organisation, helping to coordinate reporting and evaluation activities, and for securing new funding opportunities
5. A local authority representative from the education services division who had an interest in the role this programme could play in local schools.

The workshop was facilitated by 2 PhD students, including myself. A pre-defined agenda (Table 7) was agreed in prior communication with the programme team, providing the structure for the workshop. The agenda sought to divide the day up into several sessions: the aim of this was to allow adequate time to create the map as well as adequate time for reflection and conversation on specific evaluation issues facing the organisation. The overall aim of the workshop was to begin to think about and develop the programme theory as a starting point for developing more concrete evaluation plans.

Time	Session focus
11.00 – 11.10	Scene setting: What's the problem
11.10 – 11.30	Introduction to today's session: What will we be doing?
11.30 – 12.30	Brain storming outcomes for the programme: what are we aiming to achieve?
12.30 – 13.00	Delivery model: how does what we do help achieve these outcomes?
13.00 – 13.30	Lunch
13.30 – 14.30	Evaluating: how can we build evaluation into the delivery model?
14.45 - finish	Session summary: what we have achieved, and what are the next steps, looking forward?

Table 7 Organisation 2 OMT workshop agenda

As a result, the day was structured as follows:

1. The first item on the agenda focussed on setting the scene for the day. I led this session and addressed questions such as why we were holding the workshop and what we were going to do. I also provided an introduction to the Oval Mapping Technique which we were to use in the following sessions. This involved explaining the mapping methodology as well as practical details about how to use sticky notes and what kinds of things to write on them.
2. The start of the day focussed on brainstorming and mapping the key outcomes which this group of stakeholders perceived the programme to address, including the organisation and policy objectives they sought to influence at a wider level. During this session we made sure to cluster the outcomes iteratively as they related to one-another. Sometimes this was because multiple participating stakeholders had written the same thing on sticky notes or because there were related ideas. At this point, we did not remove any sticky notes from the wall.
3. A period of time was then given to thinking about the delivery model of the programme and how the programme contributes to changes in outcomes. Most of the time during this session was spent thinking about the key components of programme delivery and the implementation of these components. Naturally, participating stakeholders were concerned about how this linked to the outcomes mapped in the

previous session and so some time was afforded in this session to link up the ideas. At the end of this session, we had a busy map which was to serve as the basis for the sessions after lunch.

4. After lunch, we focussed on discussing the emergent properties and content of the map and the significance of this in terms of programme evaluation. This involved looking at “busy” concepts, particularly those outcomes which were central in the map. The outcome of this session was a prioritised list of 6 programme outcomes which allowed us as a group to start thinking about what this mean for programme evaluation activities.
5. A portion of time was allocated at the end of the workshop to sum up learning and feedback on how stakeholders felt about the benefit of the workshop. This proved a useful time for the participating stakeholders to think about and reflect upon some of the key learning points from the workshop and for me to reflect on what, if any, was the value of developing the programme theory in this case.
6. A final stage of this work involved my writing up a short report on the workshop included key learning points from the day and a computerised version of the map created. The intention of providing a written report was that the organisation would have a tangible output from the workshop that could be used in the future. I sought feedback on the written report to ensure it summarised the day accurately. I also collected additional feedback on the workshop from participating stakeholders to amplify my own reflections on the day.

Following the agenda outlined above, the workshop involved the participating stakeholders collectively developing the programme theory using the OMT principles and approach. They did this by articulating their ideas, writing these on sticky notes, and using a blank wall to stick up, discuss, structure, and link the ideas and thoughts, where relevant. Post-workshop, I used *Decision Explorer* software to transform the map, as best I could, into a computerised version that the organisation could keep as a tangible output of the workshop,

alongside the report. The map made on the day proved impractical to keep due to its size and messiness. In Images 1-3, we can see images from both the workshop and, in Figure 13, the computerised model which was delivered as an output of the workshop.

Images 1, 2, and 3 are photographs of the OMT workshop in progress. The serve to illustrate the nature of the workshops where the sessions were hands-on and participatory. In the beginning, the sticky-notes were roughly stuck on the wall by workshop participants. By the end of the morning sessions, whilst the map was still messy, it had some structure (hierarchical and clustering) – see Image 1 and Image 2. As facilitator, I helped participants to think about clustering sticky notes into related themes and ideas. In Image 3, the role I took can be seen. Whilst I was present, the construction of the map was primarily undertaken by participating stakeholders. This process of articulating ideas onto the sticky notes, sticking them on the wall, and structuring the ideas, even if only roughly, prompted other ideas which participants could write onto the sticky notes. It was not intended that an outcome of the workshop would be a well-defined programme theory, rather that there would be better appreciation of what the programme theory is, and how this can be used to think about programme evaluation.

The computerised model of some of our work on the day is shown in Figure 13. I have only included the part of the model that focuses on outcomes, as this provided much of the focus for the workshop participant discussions. In Figure 13, there are concepts which have dotted arrows with numbers attached: these highlight ‘hidden’ concepts i.e. concepts which are in the larger map. However, because I have simplified the map to just focus on outcomes, Decision Explorer ‘keeps’ the concepts which are not in view by including them as ‘hidden’ concepts. In this case, the hidden concepts relate mainly to components of programme delivery and the barriers to achieving outcomes. I also colour coded outcomes to reflect programme-level outcomes (green),

organisational outcomes and aims (orange), and Scottish Government policy objectives (blue).

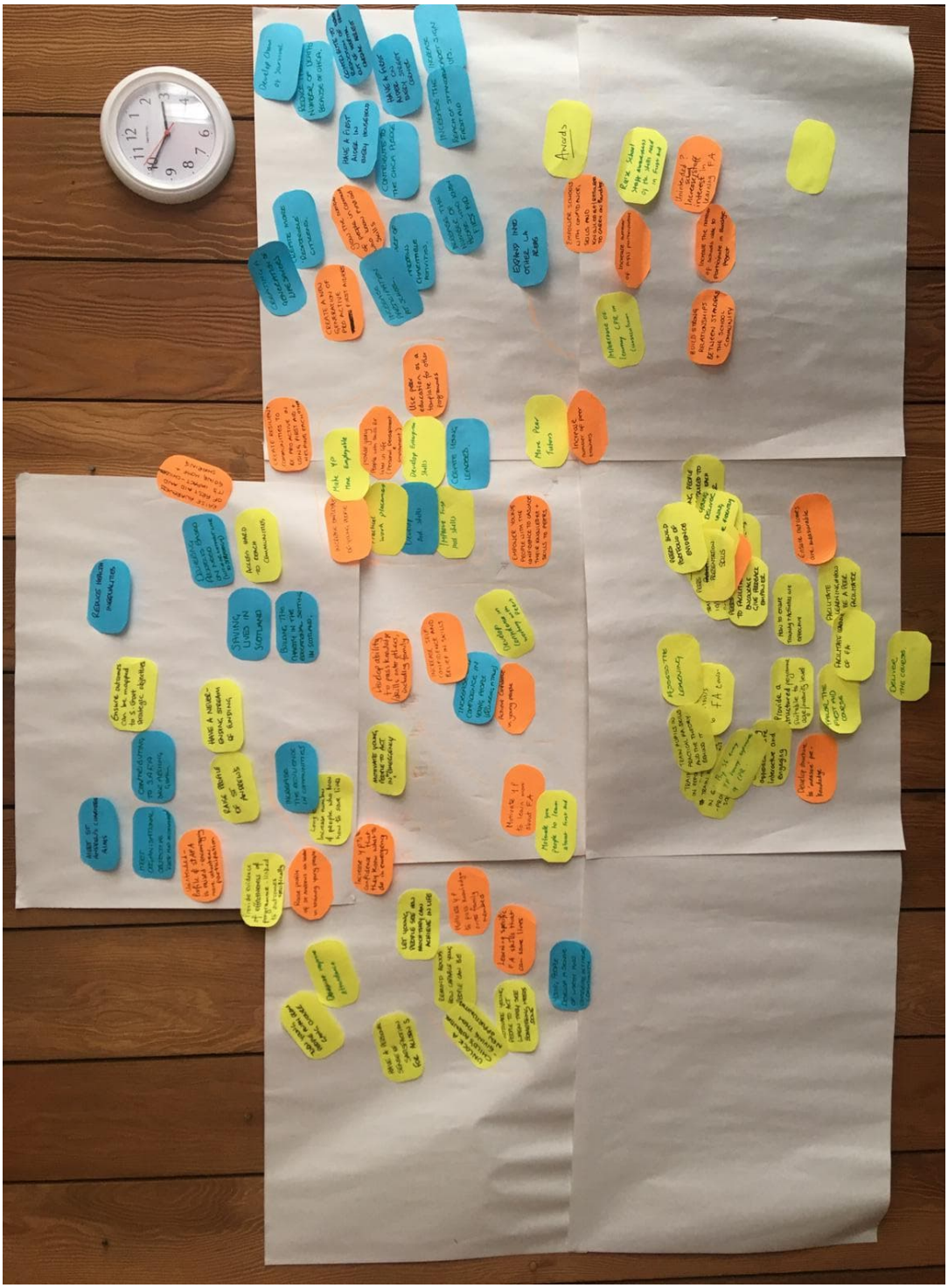


Image 1 Map at the end of the morning sessions

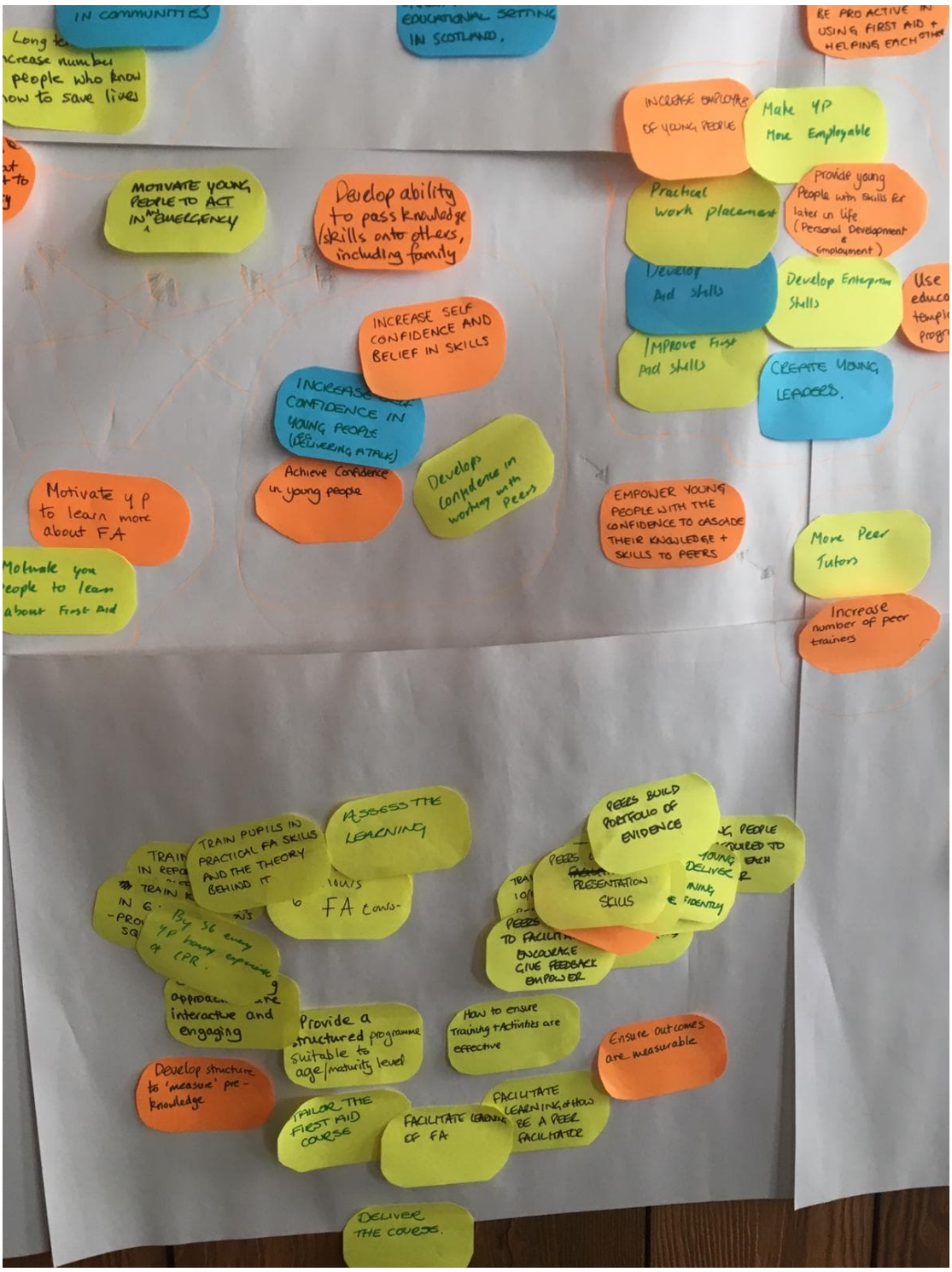


Image 2 Clusters of ideas in the map

Image 3 The mapping workshop in action - pictured (L-R): PhD student, local authority representative, training manager

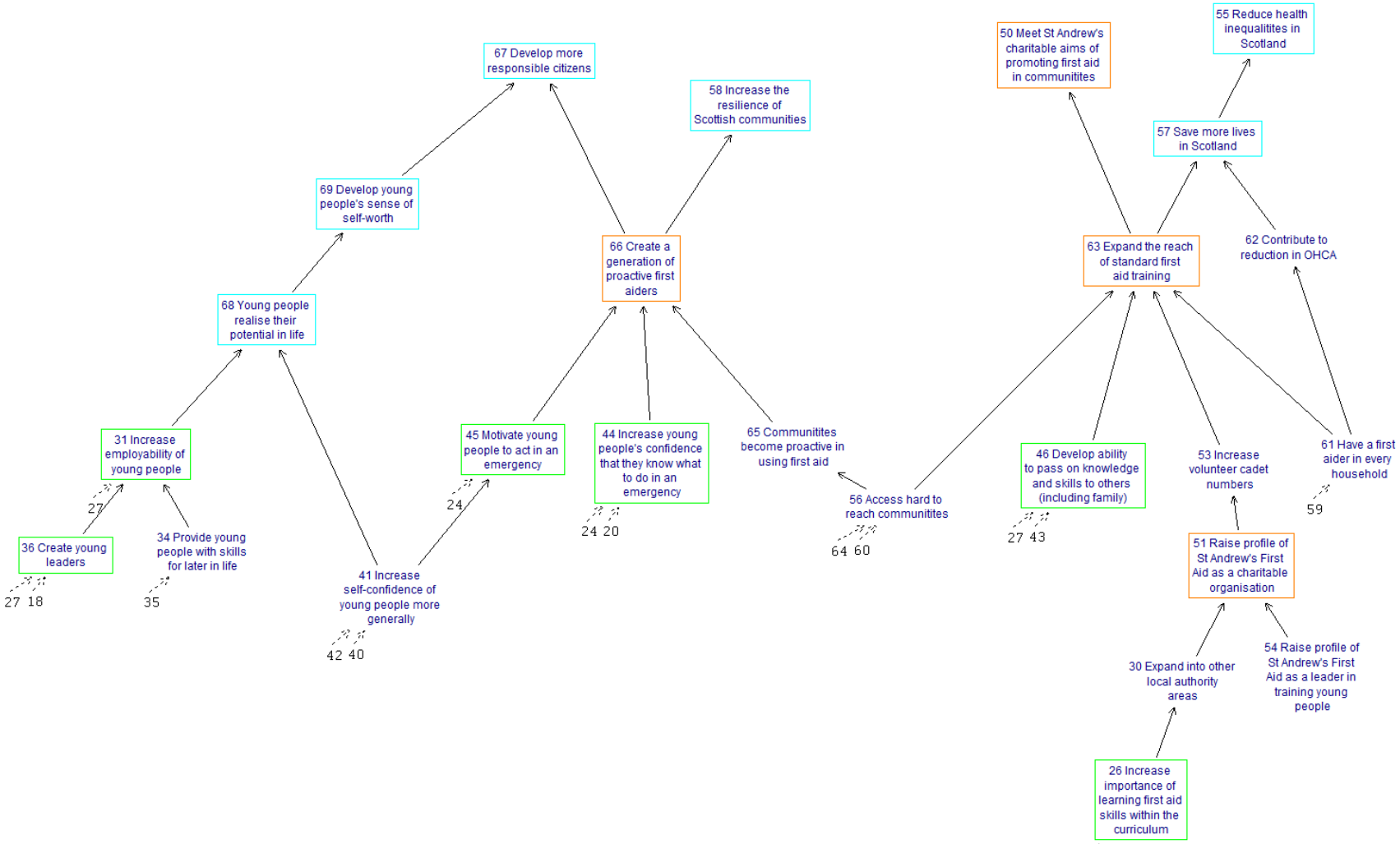


Figure 13 Outcome map emerging from larger map

5.2.2.5 Key learning

The evaluation challenges faced by Organisation 2 meant that the aim of the workshop was to help the participating stakeholders in developing the programme theory to serve as a basis for developing a framework for evaluation. The overall aim of evaluating the programme was to secure funding for the expansion of the programme into other schools. The Oval Mapping workshop involved the development of the group map. The map then stimulated some discussion around the relevant evaluation challenges facing Organisation 2. The discussion focussed on three things: the learning from developing and working with the maps; the implications for identifying, prioritising, and understanding outcomes; and the implications for conducting evaluation. I discuss these in more detail below.

Learning from developing and working with maps

There were benefits to the participating stakeholders in experiencing and interacting with the mapping tool first-hand to develop the programme theory. The visual nature of the mapping tool further facilitated these benefits. One participating stakeholder noted that, *"the tool was easy to understand once the visual effect became obvious…you need to experience it to benefit from it" (written feedback on workshop, received on 16th May 2018).* In other words, through the use of the visual mapping tool, participating stakeholders had the opportunity to have hands-on experience of the tool to develop the programme theory. As a result, they saw the value of the map emerging from the process of creating it.

By having the opportunity to participate in the mapping process and develop the programme theory as a group, participating stakeholders also had the opportunity to have joint conversations about some evaluation challenges which they faced. A core component of the Oval Mapping Technique is the ability, through the mapping process, to negotiate, discuss, and develop consensus and directions for action (Eden and Ackermann, 1998). Of note was

that the discussion stimulated by the process of mapping was not necessarily directly related to the content of the map, but rather the process highlighted some additional challenges which the organisation faced with respect to evaluation and funding. These discussions were to do with being able to better facilitate communication between members of the organisation in the future. For example, there were challenges highlighted between the grant manager and those developing and delivering the programme. The grant manager felt that having a better understanding of how the programme was being implemented (as highlighted in the content of the map) could enable her to strengthen grants applications. The result of this discussion was a plan to address this issue through arranging a programme site visit for the grants manager *(workshop notes, 17th April 2018)*. In this case, the realisation of the grant manager that she lacked understanding of programme development and delivery (covered in the mapping workshop) prompted a discussion about strategies to overcome this challenge.

It was this process, and the emerging visual content of the map and programme theory, that stimulated discussions around important ideas about outcomes and how to evaluate those outcomes. In particular, the programme theory offered a function, whereby it confirmed stakeholder understanding of how the programme was working to achieve outcomes: it was this confirmed understanding that moved discussions forward. I explain more about these ideas in the following sections.

Implications for identifying, prioritising, and understanding outcomes

The process of developing the programme theory facilitated discussion about programme outcomes. Visually mapping the programme outcomes, as well as linking the programme outcomes with how the programme contributes to achieving those outcomes, enabled the participating stakeholders to cluster and identify outcome areas of interest. In particular, the participants were able to prioritise outcomes and identify which outcomes were short-term and would

in time lead to more intermediate outcomes. This led to a list of 6 prioritised outcomes which is shown in the picture in Image 4. Alongside the 6 prioritised outcomes, Image 4 shows some sticky notes showing where the participating stakeholders had started to think about how to go about evaluating these programme outcomes.

Towards the end of the workshop, participating stakeholders were able to spend some time thinking about how to go about evaluating the prioritised outcomes. The list of outcomes then facilitated discussion around which outcomes were realistic and necessary to capture in programme evaluation. Through the ability to identify and prioritise the outcomes in this way, shown in Image 4, participating stakeholders also noted the value in being able to confirm a sense of mutual agreement that they understood what the programme is trying to achieve and how it is doing so, as a stepping stone to thinking about outcomes. One participant stated, *"The programme theory is sound. What the tool activity did was reinforce the need to develop/strengthen the evaluation structure around it" (written feedback on workshop, received on 16th May 2018).* In confirming that there was a mutual agreement on the programme theory, participating stakeholders were able to think more concretely about what the identified and prioritised outcomes meant in terms of planning programme evaluation. In the picture shown in Image 4, participating stakeholders used spare sticky notes to begin to think about how data, already being collected by the organisation, could support the evaluation of the outcomes, and where there were gaps in their knowledge of outcomes. For example, we discussed the challenges in evaluating the outcome of 'increase young people's confidence that they know what to do in an emergency'. The participating stakeholders acknowledged that this could not be measured directly but discussed potential strategies such as 'self-report' and 'scenario-based assessment' that could be used as outcome evidencing tools.

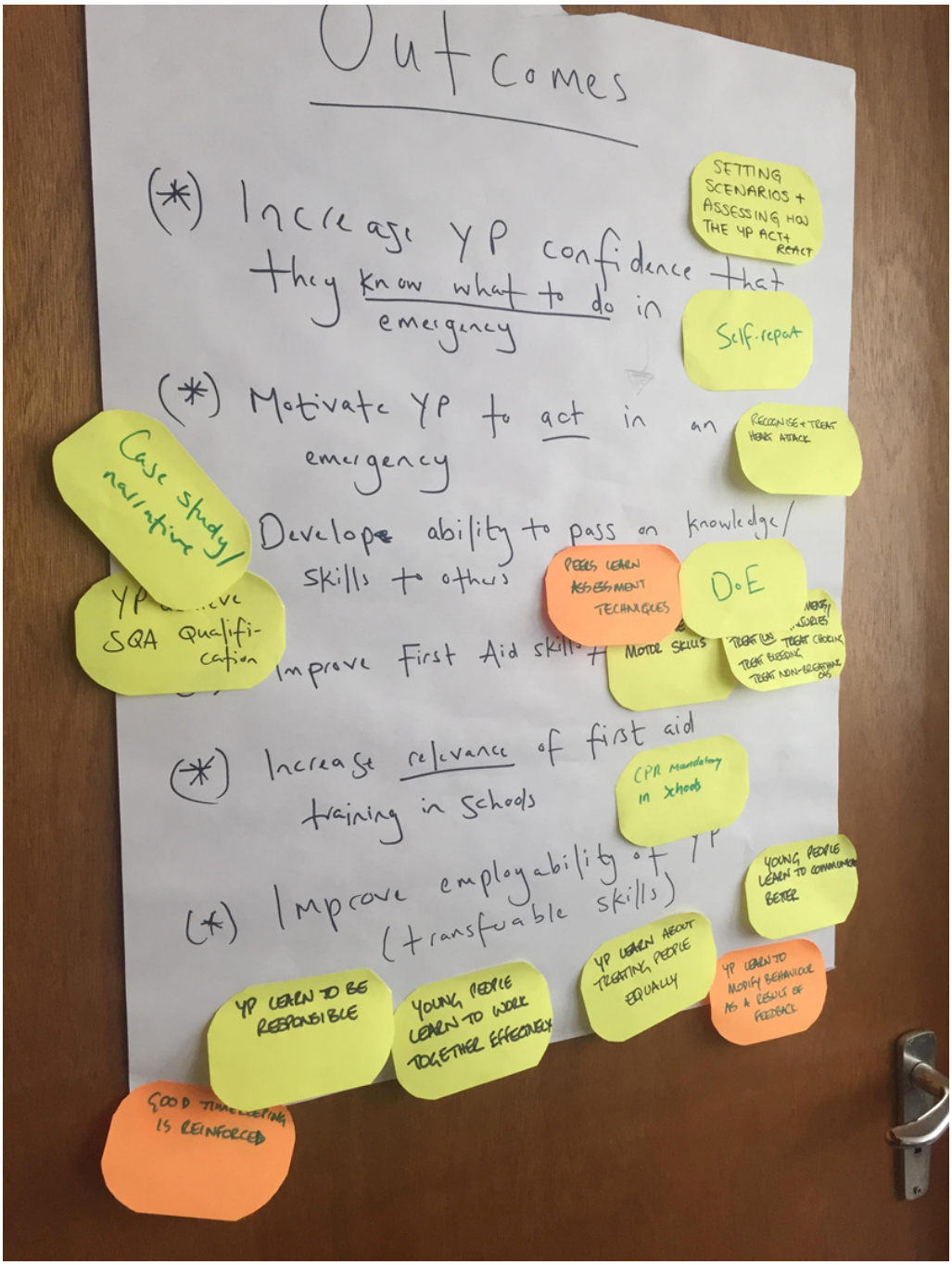


Image 4 List of prioritised outcomes and evaluation ideas emerging from OMT workshop

During the mapping process, we were also sure to address organisational and policy objectives in the session on outcome mapping. In this respect, another area for concern for this organisation at the outset was the alignment between programme outcomes, organisational objectives and Scottish Government

policy objectives. The stakeholders sought to understand better how programme level activity aligned with organisational objectives and governmental policy objectives, in order to be able to demonstrate policy relevance whilst also maintaining a clear organisational mission. Ensuring this was considered vital, when applying for funding, so that a case could be made for the relevance of the programme at the policy level, rather than just contributing to individual changes at the programme level. One interesting aspect of the workshop was the emerging acknowledgement of a lack of clear linkage between the programme outcomes (e.g. improve employability of young people) and career-oriented policy-objectives. This stimulated some discussion on better defining the links between the different levels of outcomes.

By considering the wider impact of the programme, at both the organisational and policy levels, participating stakeholders commented on how this understanding enabled them to think more purposefully about outcomes, not only at the individual programme level, but also across the implemented, multiple community-based programmes. One participating stakeholder commented, "*The importance of [understanding programme] specific objectives tying in with strategic objectives is that there are probably 3 key objectives that could be the foundation of every community programme, and that would be unique to us. Beyond that, there could be one or two objectives that would be tailored to the participant group" (verbal feedback in workshop, 17th April 2018).* Being able to consider a wider range of outcomes in this way enabled participating stakeholders to think more holistically about how the programme fits into the organisational mission and vision as well as within the wider policy context. As such, this wider consideration of outcomes enabled participating stakeholders to think about programme evaluation frameworks at the organisational level.

Implications for conducting evaluation

As I have already indicated, the process of identifying, prioritising, and better understanding the programme outcomes proved valuable to thinking about evaluation. By being able to identify the key outcomes (shown in Image 4), as well as understanding how these key outcomes linked with the delivery model of the programme, participating stakeholders were more confident in the programme theory and thus able to think more clearly about how to go about evaluating programme implementation and outcomes. Participants acknowledged that it was easier to think more methodically about how to go about measuring outcomes in the knowledge that the programme theory is clear. One participant stated, *"This is the first time that I have been able to see the clear pathways of what we do" (verbal feedback in workshop, 17th April 2018).* It seems as though being more confident about the links between programme delivery, change, and programme outcomes, and seeing these links in a visual way, enabled participants to think more methodically about how to go about evaluating outcomes.

Summary of findings from Organisation 2

In Organisation 2, we developed programme theory due to the need to evaluate the programme outcomes for future funding applications and partnership development. At the point of conducting the workshop, there was a lack of planning about how to conduct the evaluation. Using the Oval Mapping Technique, a group of programme stakeholders collectively developed a map of programme with the support of 2 PhD students. The process of developing the programme theory from scratch, in a visual and participatory way, helped participating stakeholders confirm their understanding of how the programme is working to achieve change in targeted outcomes. Confirming this understanding in a visual way facilitated more purposeful thinking about how to plan evaluation around the prioritised outcomes. Moreover, by taking into account the alignment of programme outcomes with both organisational and policy outcomes, participating

stakeholders could think more critically about how the evaluation efforts across the organisation could be more improved and more consistent. Finally, having the time to discuss the emerging content of the map highlighted some challenges facing the participating stakeholders which were not obvious from the outset. This facilitated some thinking around potential strategies to address such challenges and to strengthen evaluation processes for future funding applications.

5.3 Comparing the applications and summarising key learning across the organisations

The preceding sections that describe the application of the SODA methodology to develop programme theory in each organisation as well as the respective learning emerging from each application are summarised in Table 8: this table summarises the preceding information in terms of the reason for developing programme theory, previous programme evaluation efforts, the approach taken to develop programme theory, and the key learning in each organisation. The table serves to highlight both the difference between the applications and the learning generated from each application. I now use this comparison to consider common points of learning across the applications.

	Organisation 1	**Organisation 2**
Reason for developing programme theory	Required better understanding of how and why the programme was working for participants ahead of rapid expansion	Required stronger framework to support more systematic evaluation of social impact to apply for further programme funding and develop new partnerships
Previous programme evaluation	Logic model from programme planning phase; internal evaluations	Draft evaluation plan (including broad evaluation questions)

Approach to developing programme theory using SODA methodology	Cognitive mapping and facilitated discussion	Oval Mapping Technique in workshop setting
Key learning	Illuminating challenges which had previously been acknowledged but not addressed Considerations of strategic actions around programme implementation Considerations of improvements in outcome evaluation processes	Visually seeing pathways Highlight new challenges Identifying, prioritising, and understanding outcomes Confidence to evaluate the right things Considering more consistent evaluation across the organisation
Role of developing programme theory in key learning	Visual map Confirmatory function	Visual map Confirmatory function Participation & discussion

Table 8 Comparing the application of the SODA methodology and the emergent learning in each organisation

There are several interesting learning points that have emerged regarding *how* the development of programme theory could improve programme evaluation practice in small TSOs. From the findings presented above, there are a two common learning points with respect to *how* developing programme theory in third sector organisations can help improve future programme evaluation practice. These include the confirmatory function of developing programme theory; and working with visual and tangible maps.

The confirmatory function of developing programme theory

In both organisations, participating stakeholders noted that the emerging content of the mapping exercises were generally unsurprising, or that it confirmed what they already knew about how the programme was working to achieve the targeted outcomes. Nevertheless, participating stakeholders in each organisation commented on the value of seeing familiar information captured in a more systematic, detailed, and visual way.

In Organisation 1, making explicit and visualising, familiar information in a new way highlighted challenges already acknowledged by stakeholders. In doing so, stakeholders were able to think about those familiar challenges in a new light, stimulating more critical consideration of strategic actions to address these challenges. In Organisation 2, the making explicit of familiar information in this way served a confirmatory function. By developing and visualising the programme theory, participating stakeholders established that they all agreed and understood how the programme was contributing to outcomes. In this knowledge, stakeholders then felt comfortable and confident, thinking more concretely about how they could go about measuring outcomes and evaluating the programme.

Overall, through serving this confirmatory function, and systematically capturing and presenting familiar information in an explicit and visual way, the development of programme theory seemed to stimulate thinking around two important aspects of programme evaluation simultaneously; first, about outcomes and outcome evaluation and second, about programme implementation. Through making what was already known more explicit, it seems that participating stakeholders, in both cases, were able to think more purposefully and critically about programme implementation, outcomes, and outcome evaluation.

Working with visual and tangible maps

Whilst the SODA methodology was applied differently in each organisation, participating stakeholders in both organisations commented on the fact that experiencing the visual nature of the maps was central to realising the learning emerging from the process of developing and discussing the content of the maps. In Organisation 2, this was certainly clear due to the fact that the OMT workshop involved the participating stakeholders using the SODA mapping tool to develop their own programme theory and to literally be able to *see* the map emerging during the workshop. In this organisation for example, there

was value in being able to see the clear pathways of how programme delivery links to organisational and policy level objectives and thus work with tangible information. In the case of Organisation 1, the value was realised through the ability to interact and explore the visual and tangible maps. The visual nature of the maps in this case enabled participants to *see* challenges they knew existed in a different light (due to the detailed nature of the map, the visual links, and the use of young peoples' language) which in turn helped them question why they had not done anything to deal with those challenges. In turn, participants could begin to think about which strategies might help meet the challenges. Overall, across both cases, the visual nature of the map enabled participants to think more purposefully about implementation and outcome evaluation through illuminating, in different ways, previously acknowledged challenges in different ways (as in Organisation 1), and through visually showing and confirming the links between outcomes at multiple levels (as in Organisation 2).

Reflexivity and positionality

As a final point in this chapter, I wish to consider an important aspect of conducting action research: the role and involvement of the researcher. I seek to comment on my own involvement within each organisation to put the emergent learning into context with respect to the extent of my involvement in each process. Commenting on my involvement in this way addresses the issue of reflexivity (Robertson, 2000) in action research: this means considering my self-awareness, not only about my influence on the research process itself, but my self-awareness in how we arrived at the learning presented in this chapter came about. Robertson points to 2 other important aspects of action research which I have already addressed in both the methodological framework (chapter 3) and in this chapter (chapter 5). These other components are reflection and reciprocity.

First, it is worth noting that I worked in a slightly different way with each organisation. Despite utilising the same overall methodology in each organisation, and thereby standardising to some extent the tool they used, it was applied differently in each case. With Organisation 1, I played a more central role in developing the maps given that I conducted the interviews with young people and built the maps in the first instance. In terms of the facilitated discussion, I simply provided information on some of the key learning points emerging from the content and analysis of the map; the rest of the time was spent exploring, discussing, and questioning the content of the maps, led by programme management. Nevertheless, programme staff and management were also involved at every stage, even if only indirectly. I was also involved with Organisation 1 over a long period of time and had the opportunity to get involved in meetings and events organised by the programme team. Whilst not directly related to developing the programme theory, this gave me a good understanding of the organisation and the challenges they were facing in terms of evaluation. This also enabled programme staff and managements to become more familiar with me and hopefully feel more comfortable in reflecting openly and honestly about their thoughts on the content of the maps, and on the usefulness of this information.

In Organisation 2, I played a slightly different role. Things happened over a much shorter space of time (4 months) as opposed to over a year with Organisation 1. There was therefore less time to be as involved as I was with Organisation 1 in terms of getting to talk to programme staff or attending events. However, in proposing an OMT workshop, there was direct involvement from the programme stakeholders throughout the whole process: from the very first meeting right through until the workshop. Together, we were able to identify needs, develop a workshop proposal, and conduct an OMT workshop. In this sense, the work with Organisation 2 was wholly collaborative, even if I was not as involved with the organisation over the same length of time.

A related point is the role I played in both the facilitated discussion and the workshop. Certainly, in both organisations, I had to inform them about the principles, methodology, and process of mapping and how to interpret the emerging content. In Organisation 1, this involved a brief explanation at the beginning of the facilitated discussion. Of course, the programme management team were already aware of the approach from when the study was proposed. Similarly, in Organisation 2, I spent some time at the beginning of the workshop explaining the principles of the approach and how we were to use it. The difference between the 2 organisations was in the extent to which the participating stakeholders had hands-on experience on the SODA methodology. Whilst Organisation 1 had the opportunity to explore already created maps, Organisation 2 used the SODA methodology to build their own collaborative map. The reason however for these different approaches was because of the evaluation specific challenges faced by each organisation.

Overall, my final point to consider, with respect to reflexivity, is the enhanced capacity afforded to the organisations through my involvement in developing the programme theory with them. A question is whether the organisations would have been able to develop programme theory had I not been involved. I imagine they would not have done so but that does not mean they were not able. Whilst I provided a resource to facilitate the development of programme theory, I believe that both organisations have the skills and knowledge about how their programmes work to develop programme theory. What I do think I offered the organisations was the space and time to develop and explore the programme theory, as well as a structured tool and process by which to do so. With respect to enhanced capacity, it is also plausible to think that those programme stakeholders involved may have perceived my involvement, as a PhD student aligned with an academic institution, to enhance organisational legitimacy: this may have been through ensuring that the process of

developing programme theory were facilitated by someone with experience with this evaluation theory.

CHAPTER 6 DISCUSSION

6.1 Introduction

In the preceding two chapters I present the findings of the empirical component of this thesis. I now move to discussing these findings with respect to how they address the two research questions defined in Chapter 2, which are:

> RQ1: What is the current practice with respect to the use of programme theory in programme evaluation practice in small TSOs?
>
> a. What is the context of programme evaluation practice in small TSOs with respect to the capacity to do evaluation, evaluation activities, and the use of evaluation?
> b. In what forms, does the use of programme theory manifest itself within the context of programme evaluation practice in small TSOs?
>
> RQ2: How can the development of programme theory improve future programme evaluation practices within small TSOs (particularly with respect to facilitating evaluative thinking)?

6.1.1 Structure of discussion

The structure of this chapter is as follows. To begin, I provide a summary of the key research findings from both Chapter 4 and Chapter 5, which serve as the basis for this discussion chapter. Overall, there are several key observations I wish to draw on to discuss the use of programme theory in small TSOs, both in terms of current practice, and in terms of improving future programme evaluation practice. I then move to discuss these findings with respect to the two research questions. Where relevant in the discussion, I draw on both the extant literature, and introduce additional literature, to help explain my observations. In particular, I discuss how the current use of programme theory in small TSOs is tacit in nature, but also that its use is also impacted by other contextual factors. I then discuss how the development of programme theory can improve future programme evaluation practice in small TSOs, with particular note of how its development can facilitate evaluative thinking. In this

respect, I conceptualise the development of programme theory in terms of its role as a 'tool for evaluative thinking'. Consequently, I draw from literature on the role of modelling in organisations to help illuminate and conceptualise some of the learning around the second research question. In drawing this chapter together at the end, I present a conceptual framework of evaluation practice: this aims to capture the role of programme theory in small TSOs from a practitioner-oriented perspective. This is then used as the basis for the implications presented in the concluding chapter of this thesis.

6.1.2 Summary of key findings

Regarding the first research question, on current practice of the use of programme theory in programme evaluation practice in small TSOs, there are two key findings. First, I found that those practising evaluation in small TSOs have a good understanding of programme theory, such that their understanding, of how the programmes and services they deliver contribute to change, is clear when they talk about the work that they do. A number of reasons, relating to the small size of the organisation and professional practice backgrounds of participants, seem to suggest a good reason for this. Second, I found that programme theory was not generally used, in an explicit form, in systematic evaluation activities. Nevertheless, I find that, those who practise evaluation in small TSOs, tend to use programme theory in less systematic evaluation activity happening in an informal manner, on a day-to-day basis. A number of reasons emerged for these findings including: accidental evaluators in small TSOs do not necessarily consider less systematic evaluation activity, happening on a more informal basis, to be programme evaluation; funders do not explicitly ask for information such as programme theory; funders struggle to ask for such information because they are wary of maintaining a balance between being prescriptive and flexible in terms of reporting formats; the capacity and resource constraints in small TSOs; and the nature of the professional backgrounds of evaluators supports implicit reflection on programme theory.

Regarding the second research question, on how developing programme theory can improve future programme evaluation practice in small TSOs, there are also two key learning points: these learning points relate to how the process of developing programme theory, and the emerging content of the programme theory, can improve future programme evaluation practice (particularly with respect to facilitating evaluative thinking). The findings presented in Chapter 5 suggest that the confirmatory function of developing programme theory, as well as the visual nature of an explicit model of programme theory, were the primary means of facilitating more constructive and purposeful discussion and thinking about programme implementation and programme outcome evaluation. In what follows, I take each research question in turn to discuss the emerging findings and learning, with respect to the extant literature and, where relevant, drawing from additional literature to help explain emerging findings.

6.2 Current practice in the use of programme theory in small TSOs

The literature on the use of programme theory in programme evaluation practice conceptualises evaluation practice quite differently from how we understand programme evaluation practice in small TSOs. As a result, Study 1 sought to explore the current use of programme theory in programme evaluation practice in small TSOs with a view to better understanding and conceptualising its use. I find that the use of programme theory, in the context of programme evaluation in small TSOs, is somewhat tacit in nature and consequently, programme theory is used in 'unsystematic' programme evaluation activities occurring on a more informal day-to-day basis. Paired with its tacit nature, the lack of use of programme theory in systematic programme evaluation activities is also influenced by resource constraints and funder reporting requirements. All this comes in spite of a desire, from both small TSO and funder perspective, to use programme theory in more explicit and systematic ways in programme evaluation.

In what follows, I first discuss the practical, and contextual, influences on the use of programme theory in small TSOs. I then discuss how the current use of programme theory could be defined as tacit in nature, which has implications for programme evaluation, and can also restricts the use of programme in systematic programme evaluation activities.

6.2.1 Practical and contextual influences on the use of programme theory in small TSOs

Findings emerging from this study, which focuses specifically on the use of programme theory, are consistent with the wider literature on programme evaluation practice in small TSOs, such that programme evaluation efforts are driven by accountability demands from funders, and that time, resource and expertise to conduct programme evaluation are also noted as constraining factors (Bach-Mortensen and Montgomery, 2018). Whilst I find that there is a desire, from both the funder and TSO perspective, to use programme theory more explicitly, the findings of this research also indicate that there are other, unsurprising, practical challenges which might influence the explicit use of programme theory in systematic programme evaluation activities, namely resource limitations and funder expectations.

With respect to the use of programme theory, the findings presented in Chapter 4 indicate that there is an important role for the funder (i.e. those who are requesting evaluation reports) in influencing programme evaluation practice in small TSOs. When asked what determines the types of information and data included in formal (systematic) evaluation activities and subsequent reports, many of the participants indicated that it depends on what the funder asks for in funder-specific reporting mechanisms. This reflects, what is described as, the linear nature of knowledge transfer, which characterises the third sector more generally: knowledge, in this case formal (systematic)

programme evaluation, moves only in one direction in the form of packaged information from producer (small TSO) to user (funder) (Hardwick, 2018). The consequence of this is that funders, or, in some cases, commissioners, of third sector services, may be missing out on the very distinctive knowledge that small TSOs possess, and that which characterises the value of small TSOs (Hardwick, 2018).

However, funders did indicate that, the kind learning relating to programme theory, particularly the understanding and learning about how change occurs, was of interest to them. This desire was driven by a need from funders to better support learning about social change strategies, particularly with respect to social change within the communities where funded work was being undertaken by TSOs. Participants from funding organisations commented on how the use of such information could improve and support 'strategic funding', through better understanding the social problems small TSOs are tackling, and the strategies used to address them.

Despite this, there was an acknowledgement, from participants in funding organisations, that they face a challenge in terms of defining the format of reporting requirements such that the transfer of knowledge from TSOs to funders could stimulate action. However, participants found that it was difficult to find a balance between defining reporting mechanisms, which are prescriptive of the information that funders desire, but at the same time, defining mechanisms that afford flexibility to small TSOs to be able to include the information and data about their programme that they see as relevant and useful. From both the funder and small TSO perspective, it was clear that reporting mechanisms needed also to be proportionate to the capacity of, and need for, small TSOs to evaluate and report on their work.

It may be that there is a role for funders to develop reporting mechanisms that generate this kind of learning, without being overly prescriptive about how small TSOs go about generating such information. In such instances, researchers argue that this gap is a 'knowledge transfer challenge', and that approaches to knowledge generation could be driven more by co-production and engagement, supporting more inclusive and collaborative knowledge generation efforts (Van de Ven and Johnson, 2006). A new approach to knowledge transfer would stand in contrast to systems that simply support a linear transfer of knowledge. It is therefore worth considering whether defining new reporting mechanisms would solve this knowledge transfer issues, or whether more systemic issues regarding knowledge use in the sector need to be challenged.

6.2.2 Defining the use of programme theory as tacit in nature

Alongside the practical and contextual influences on the use of programme theory in small TSOs, there was another emerging influence on its use: this was the use of programme theory as tacit in nature. There are several reasons why I conceptualise current practice in the use of programme theory, in the context of programme evaluation in small TSOs, as tacit in nature. But first, let me define what I mean by tacit in nature. The field of knowledge, including tacit knowledge, is wide in scope with many different definitions of tacit knowledge, as well as disagreements over what constitutes such knowledge (McAdam, Mason and McCrory, 2007). It is not within the scope of this chapter to discuss and debate the entirety of this literature. However, in what follows, I define some key concepts, within the field of tacit knowledge, that can help to better understand the emerging findings of this study.

I interpret tacit knowledge to be a dimension of all personal knowledge relating to things that we know, but that we can find it difficult to articulate (Polanyi, 1962, p.601). These things that we often find difficult to articulate commonly relate to some form of 'know-how' rather than 'know-that', such that we can do

things and perform activities without necessarily being able to articulate exactly what it is that we are doing. In this way, Polanyi states that tacit knowledge forms a necessary ingredient of all knowledge which can be considered as the 'bottom of the iceberg' for all that people know (Polanyi, 1966; Pyrko, Dörfler and Eden, 2017). Definitions of tacit knowledge emphasise that because knowledge is tacit, that it can be difficult to make explicit, or articulate. An example of tacit knowledge might be the knowledge required to know *how* to swim. It would be incredibly difficult to tell someone exactly how you swim but this tacit knowledge allows you to swim when you get into a swimming pool, even if you cannot tell someone else exactly how you do it. Using the 'bottom of the iceberg' analogy, I interpret tacit knowledge as one point on a continuum, rather than as a distinct type of knowledge. In this sense, just because tacit knowledge is difficult to make explicit, or articulate, does not mean that it reflects something that is wholly untellable: rather, the tacit-explicit (know-how, know-that) knowledge division has been conceptualised as a continuum, rather than a dichotomy. As such, the tacit-explicit distinction reflects different dimensions of knowledge along a continuum, rather than different or distinct types of knowledge per se.

Placed on a continuum, tacit knowledge can also be embedded in experience, where the sharing of 'know-how' is based on 'knowledge-in-practice' (McAdam, Mason and McCrory, 2007). Drawing on previous literature on tacit knowledge, McAdam and colleagues define 'knowledge-in-practice' as that which is "*developed from direct experience and action; highly pragmatic and situation specific; subconsciously understood and applied; difficult to articulate;* [and] *usually shared through interactive conversation and shared experience*" (2007, p.46). In the context of TSOs, such knowledge, more generally, has been shown to be used to support programme delivery to programme beneficiaries (Hardwick, 2018), rather than used within systematic evaluation activities for funders. Likewise, I find that the use of programme theory is often developed from direct experience with the programme, is well understood, is highly contextual, and usually shared through unsystematic

evaluation activity informing the day-to-day practices in small TSOs. I discuss this conceptualisation in more detail in the following section (6.2.3)

Nevertheless, defining the understanding of programme theory as tacit in nature is not necessarily new. In the literature on developing and using programme theory in practice, the focus is often on the making explicit of the programme theory through eliciting the 'mental models' of programme stakeholders and beneficiaries. That is not to say that it is tacit per se, but rather that it is knowledge which has not yet been made explicit, but nonetheless that which can often be challenging to elicit, according to some academics (Friedman, 2001). The case of 'accidental evaluators' presents an additional challenge in this respect. As with much of the literature on programme theory, the focus is on working *with* programme stakeholders do develop and use programme theory. Therefore, it is the task of some internal or external evaluator to elicit the implicit understandings of someone else, e.g. programme stakeholders, rather than making one's own, and others', mental models explicit. Consequently, in the case of accidental evaluators, it could be such that there are challenges to undertaking the task of making programme theory explicit, when accidental evaluators are the ones conducting the evaluation activities alongside the multiple other roles they serve.

6.2.3 The use of programme theory in small TSOs as tacit in nature

There are several reasons why I define the use of programme theory in small TSOs as tacit in nature, rather than say that which has simply not yet been made explicit. The concepts within the field of tacit knowledge, which I have given a brief overview of above, can serve as the basis to consider the use of programme theory as tacit in nature. In particular, the findings, presented in Chapter 4, suggest that those practising programme evaluation in small TSOs have good understanding of programme theory and social change, which they make use of in 'unsystematic' programme evaluation activities that inform day-to-day practices in small TSOs. I find that this is probably a consequence of

several factors: namely, the closeness of those conducting programme evaluation to the programme beneficiaries with whom they work; the professional training backgrounds of those practising programme evaluation; and the organisational support structures that support communication, honesty, and trust between and amongst programme staff. I now discuss these findings in terms of the evidence supporting this assertion. I then discuss the explanations why this might be so followed by a consideration of some of the consequences of programme theory being tacit in nature.

6.2.3.1 The understanding of programme theory and social change

The findings, presented in Chapter 4, indicate that those practising programme evaluation in small TSOs have a strong understanding of programme theory, as exhibited in how they talk about the work that they do and the programmes they deliver. These findings highlight that those practising evaluation in small TSOs do not tend to make use of programme theory in an explicit format in systematic programme evaluation activities (e.g. through the use of a model of programme theory). However, I find that those practising evaluation in small TSOs tend to have a strong understanding of how the programme is working to achieve change, particularly in how they talk about both the programme and the work that they do, as an organisation, more broadly.

These findings suggest that this strong implicit understanding of programme theory could be a consequence of two things, namely the closeness of those practising evaluation to beneficiaries and programme delivery, and their professional training backgrounds e.g. training in social work or youth work. Often what has previously be considered to negatively affect capacity to do programme evaluation (e.g. small size and being an accidental evaluator) (Rog, 2015), in the case of these findings, suggests an increased capacity to at least understand programme theory, and thus, how change occurs for programme beneficiaries. Overall, these findings, regarding the understanding of programme theory, are consistent with research on knowledge use in the

third sector more generally (introduced in Chapter 1). This work suggests that there is strong acceptance, in programme and service delivery communities, that third sector organisations are 'epistemologically different' from other public sector organisations because they know the individuals, groups, and communities that they work with better, and as such, are better placed to provide services and programmes that will address the social issues these groups face (Lang and Hardwick, 2016; Hardwick, 2018).

6.2.3.2 The use of programme theory in 'unsystematic' programme evaluation activities

Additional to a strong understanding of programme theory and social change, there was evidence of the use of this understanding in 'unsystematic' evaluation activities. These findings indicate that, whilst not necessarily considered as 'programme evaluation' by participants in small TSOs, there were many opportunities in these organisations to share and discuss what was happening on the ground, particularly in terms of how programmes or services were working, or not, i.e. using the understanding of programme theory. This was mainly used to think about how services could be improved or adapted to meet the needs of the target population on a day-to-day basis. Programme theory was also used at the level of individuals to understand their needs and to ensure the programme could help that individual. Taken together, the use of programme theory in unsystematic programme evaluation activities means that it was not necessarily used to systematically evaluate programmes.

Consequently, the use of programme theory in 'unsystematic' programme evaluation activities does not meet the definition of 'formal evaluation activities' that I presented in Chapter 2, i.e. those which are 'systematic' in nature and which make use of programme theory explicitly, to guide and structure the programme evaluation. Whilst the use of programme theory in 'unsystematic' programme evaluation activities means that evaluation activities are not designed and conducted to evaluate the programme as a whole, that is not to

say the use of programme theory does not have value in 'unsystematic' evaluation activities. For example, in the findings presented in Chapter 4, participants reflected favourably on being able to share experiences of the programmes and services they deliver, in particular, to learn from each other and improve the day-to-day programme or service delivery.

Another example of the use of programme theory in 'unsystematic' evaluation activities was its use to support service delivery at the level of the individual. The findings, presented in chapter 4, also suggest that participants integrate 'individual programme theory' within service delivery: in this sense, participants used their understanding of how the programme worked to help individuals with whom the programme was working. Similarly, Hardwick (2018) finds that such tacit, or experiential, knowledge is often used by staff in TSOs to support beneficiaries and users of services to develop knowledge of 'what works for me', i.e. individuals, to guide their work more generally. In such cases, Hardwick finds that staff prefer to use this kind of beneficiary-based and experiential knowledge, as it embodies the values inherent in third sector work, e.g. user-centred provision of programmes and services.

6.2.4 Why is the use of programme theory in small TSOs tacit in nature?

Having discussed how I have concluded that the use of programme theory is tacit in nature, I now present three reasons why I believe this to be so. These reasons relate to the size of the organisations, and the subsequent closeness to beneficiaries of those practising programme evaluation; the professional training backgrounds of those conducting programme evaluation in small TSOS; and organisational culture in small TSOs. These reasons give us some insight as to why the use of programme theory could be considered tacit in nature, and how this might make such knowledge difficult to articulate.

6.2.4.1 Closeness to beneficiaries

Due to the small organisational size, those practising programme evaluation in the small TSOs, as 'accidental evaluators', are also involved in the delivery, management, and administration of programmes and services. This means that they observe, experience, and act within the programme on a day-to-day basis, and as such, are closely involved in programme operations. As a result, they frequently have close contact with beneficiaries. This stands in contrast to the view of an external evaluator, who, whilst potentially involved in observing the programme over an extended period of time, is neither involved in service delivery nor privy to the day-to-day programme operations over longer periods of time. So, in some sense, the nature of 'accidental evaluators', and their embeddedness in programme implementation, service delivery, and operations, means that they have, what Rog refers to, as high substantive understanding of the subject area (2015) based on their day-to-day experiences of and action within the programme. This high substantive understanding due to direct experience of the programme relates, in part, to the definition of 'knowledge-in-action' presented earlier.

6.2.4.2 Professional training backgrounds

To further build the case for high substantive understanding of the subject area within small TSOs, 'accidental evaluators' in small TSOs commonly have professional training in fields such as social work, youth work, or community development to name a few, rather than formal programme evaluation or social science training. Such disciplines offer practical and theoretical training in how professionals can help make changes in peoples' lives through the programmes and services that they deliver. It follows that 'accidental evaluators' often have a good theoretical understanding of how change can occur and can apply this theoretical understanding into practice through their involvement in programme/service delivery. The use of 'general theoretical frameworks' by participants in this study highlights this ability to use theoretical principles to inform and guide practice.

The professional training backgrounds of those practising evaluation in small TSOs also enables them to embed the notions of reflective practice into their work. Consequently, the more 'unsystematic' programme evaluation activity happening on an informal and day-to-day basis in this sample, might simply be a consequence of the way people working in small TSOs are trained to go about their work. The findings presented in Chapter 4, suggest that participants in this sample were able to critically reflect on practices within these more 'unsystematic' evaluation activities, even if not in a formal (systematic) way. It follows that both the substantive knowledge of social change, based on the professional training backgrounds, paired with the ability to continually reflect on practice, means that the use of programme theory, by accidental evaluators, in small TSOS is more likely to be tacit in nature due to its embeddedness in the way such 'accidental evaluators' go about their daily work.

6.2.4.3 Organisational cultures

In practice, another reason for the likely tacit nature of the use of programme theory are the organisational cultures, within small TSOS, which facilitate channels of communication, trust, and honesty between staff members. Commented on by participants in this study, the organisational culture of small TSOs can, in such cases, facilitate the ability to share and discuss programme theory more informally in 'unsystematic' evaluation activities. This was done through sharing of experiences of, and understanding about, how the programme is working. The ability to share and discuss amongst the team was a result of frequent staff meetings, peer learning events, and thus the effective communication channels amongst team members. Such organisational cultures therefore facilitate the use of programme theory to inform the day-to-day operations of the programme, rather than its use in more systematic evaluation activities.

6.2.5 Consequences of the tacit nature of the use of programme theory

Overall, whilst the literature argues that evaluation capacity is 'lower' in the case of 'accidental evaluators', in terms of knowledge and skills in programme evaluation theory, I argue that the substantive knowledge of accidental evaluators (Rog, 2015), in the context of small TSOs, supports the assertion that the use of programme theory is tacit in nature. Its use, in this way, is facilitated by the small size of the organisation, the closeness to the beneficiary, the professional training backgrounds of accidental evaluators, and the perceived trustworthy, honest and open communication channels in organisational culture within small TSOs. These factors provide those practising evaluation with a sound understanding of how change can and does occur in the groups of people with whom they work, and how the delivery of programmes and services that they provide, help contribute to those changes on a day-to-day basis, as well as the ability to share this knowledge informally with other staff. This is evidenced, not only in the ways that those practising programme evaluation in small TSOs talk about the programme, but also how programme theory is put into practice in informing programme delivery on a daily basis. In illustrating this, the findings presented in this thesis suggest that 'deficit-focussed thinking', i.e. focussing on what those in small TSOs cannot do, in the literature on programme evaluation practice in TSOs, is insufficient to explain the lack of use of evaluation theory, such as programme theory, in small TSOs.

However, the findings indicate that the tacit nature of the use of programme theory, used in in 'unsystematic' evaluation activities, is not necessarily reflected in more formal (systematic) evaluation activities. What I refer to as a positive capacity in terms of the use of programme theory, i.e. 'know-how', does not necessarily translate into explicit knowledge, i.e. 'know-that', reported in systematic evaluation activities, those that are commonly used to communicate to funders. Despite this, I find that this kind of information, that is tacit in nature, is that which funders would like to learn from. Specifically, funders seek to understand better the learning process of small TSOs about

how change takes place. The disconnect between the tacit nature and explicit use of programme theory poses the questions of how can the tacit be better translated into the explicit for more systematic evaluation activities, and whether there is value in doing so for small TSOs? Moreover, this disconnect, I find, is consistent with the literature on the use of knowledge in the third sector more widely (Lang and Hardwick, 2016; Hardwick, 2018), where what is translated or used, in terms of explicit knowledge, does not reflect the wealth of knowledge used overall, including that which is tacit in nature.

In conceptualising the use of programme theory as tacit in nature, I argue that it is perhaps something which is difficult to make explicit. I argue that this is due to its embeddedness in the knowledge, experiences, and thus daily actions practices of accidental evaluators in small TSOs. Consequently, those practising evaluation in small TSOs may not even perceive this knowledge as something which can be easily captured in systematic evaluation activities, or alternatively, the kind of knowledge that they know how to capture in systematic evaluation activities. These findings are consistent with the wider literature on the use of knowledge in the third sector: given that there is a preference to utilise knowledge generated from direct experience and action in practice, rather than knowledge which is formally generated and disseminated as 'research', or as a knowledge 'product', it is imaginable that accidental evaluators in small TSOs do not necessarily perceive knowledge generated from direct experience and action to align with the kinds of knowledge that one might include in systematic programme evaluation activities.

There are, however, consequences of the tacit nature of the use of programme theory, particularly from the perspective of 'designed blindness'. Drawing from an action science perspective, designed blindness is the differences between what people intend or believe that they do (espoused theory) and what they actually do (theories-in-use) (Argyris and Schön, 1996; Friedman, 2001).

Making programme theory explicit has, in this sense, enabled the improvement of programme delivery through identifying gaps between espoused theories and theories-in-use and bringing this to the attention of stakeholders (Friedman, 2001). Not using programme theory in systematic evaluation activities may therefore not highlight, to accidental evaluators, some of the bigger challenges they face in programme delivery, e.g. gaps between what they intend to do, and what they actually do.

6.2.6 Implicit theories of the use of programme theory in small TSOs

The first half of this chapter has focussed on discussing how the empirical component of this thesis has contributed to our understanding of the current use of programme theory in programme evaluation practice in small TSOs, from a practitioner perspective. There is strong implicit understanding of programme theory in small TSOs due to their closeness to beneficiaries and the day-to-day programme operations, their professional backgrounds in the field of programme delivery, and the organisational cultures in small TSOs. What is often considered in a negative capacity with respect to programme evaluation (i.e. small size and no formal evaluation training), strengthens the understanding and use of programme theory in small TSOs. However, it seems to be that the use of programme theory in small TSOs is somewhat tacit in nature. This is because the knowledge of programme theory used, and the ways in which it is used, is generated primarily through direct experience and action and shared in unsystematic evaluation activities. Such knowledge is therefore embedded in practice and can, as a result, be difficult to articulate in conducting systematic programme evaluation.

Nevertheless, I find that the understanding of programme theory is shared within ‘unsystematic’ programme evaluation activities which inform and guide the everyday practices in small TSOs. Despite this, there are consequences to both conceptualising the implicit understanding of programme theory as tacit knowledge and the use of implicit programme theory in ‘unsystematic’

programme evaluation activities. Importantly, a lack of consideration of programme theory in formal (systematic) programme evaluation activities might contribute to 'designed blindness'. Despite the fact that there are certainly benefits to the use of programme theory in 'unsystematic' programme evaluation, there is little consideration of programme theory in a more systematic way. This might mean that there will be difficulty in identifying gaps between espoused theories and theories-in-use and ultimately, limit the overall effectiveness of programme and service delivery in contributing to social betterment.

The lack of the use of programme theory in formal (systematic) programme evaluation activities, in the context of small TSOs, may therefore be a consequence of three distinct but related challenges: accidental evaluators in small TSOs do not have the resources and/or capacity to develop programme theory; the main drivers of how 'systematic' programme evaluation activities are carried out is expressed in terms of what funders ask for; and the understanding and use of programme theory is tacit in nature and thus difficult to make explicit. It follows that we can complement this understanding with a better understanding of how making programme theory explicit can improve programme evaluation practices in small TSOs. This understanding can help to inform strategies to support the use of programme theory, including the potential benefits of doing so. Overall, what I argue, is that it is insufficient to suggest the lack of use of evaluation theory, in this case programme theory, is a consequence of the deficits of small TSOs (i.e. the lack of formal evaluation training and resources). Rather, from these findings, we can understand that evaluation theories, in this case, programme theory, may be in use, but in different ways than the extant evaluation literature might suggest.

6.3 Improving future evaluation practice in small TSOs through the development of programme theory

The literature on the use of programme theory in programme evaluation highlights its use in formal (systematic) evaluation activities. In reviews of the use of programme theory, it is also noted that it is often unclear what the added value of using programme theory was (Coryn et al., 2011). I argue that this is because we understand little about *how* the use of programme theory helps support systematic evaluation activities, particularly from a practitioner-oriented perspective. The question of *how* can be addressed through focussing on how the use of programme theory can facilitate 'evaluative thinking'. In the context of small TSOs, the consideration of evaluative thinking is important due to the need to better facilitate the capacity to do programme evaluation in contexts where those conducting programme evaluation are 'accidental evaluators. As a result, part of this thesis sought to explore *how* the development of programme theory can improve future programme evaluation practices (particularly with respect to facilitating evaluative thinking).

The findings of Study 2 suggest that, like the less purposeful and systematic literature (Huebner, 2000; Friedman, 2001; Wimbush, Montague and Mulherin, 2012), making programme theory explicit, in the form of a model of programme theory, can act as a useful 'tool for evaluative thinking'. It can help accidental evaluators think more constructively and purposefully about programme implementation and outcome evaluation, through its confirmatory function and visual nature, much like is intended in conducting evaluability assessment (Wholey, 1987; Craig and Campbell, 2015), where the goal is to understand whether a programme is defined and implemented sufficiently that is can be evaluated (Patton, 1997, 2008).

This emergent learning suggests a number of ways to think about the role of a developing a model of programme theory, particularly in considering that much of how programme theory is currently used in small TSOs is tacit in

nature. In what follows, I draw from a literature focused on the role of models, and the process of modelling, in organisations more generally, to conceptualise how the development of programme theory can improve future programme evaluation practice. I consider three different types of roles played by models and how this relates to the emergent findings of this study. I consider the role of developing programme theory as a *requisite model*, as a *transitional object*, and in terms of the *affordances* that the development of programme theory can allow. I also consider my own role in this research in terms of the importance of facilitation. The findings from this study reflect the beginnings of a novel conceptual framework to support the understanding of *how* developing programme theory can facilitate evaluative thinking, and thus support, and improve, programme evaluation practices in small TSOs.

6.3.1 Development of programme theory as a requisite model

In both cases of action research, the models were not developed to test a particular programme theory, through conducting an implementation or an outcomes evaluation, as we frequently see reported in the literature. Instead, the aim of developing programme theory was to help the organisations deal with some of the evaluation-related challenges they were facing. In this way, the models of programme theory, through their confirmatory function and visual nature, were able to stimulate thinking about these evaluation challenges.

In this way, the models of programme theory could be considered as sufficient for the purposes the organisations pursued. This is similar to the literature on 'requisite models' where, a *requisite model* is "a model whose form and content are sufficient to solve a particular problem" (Phillips, 1984, p.29). In the case of this research, the sufficiency to solve a particular problem was the evaluative thinking around programme implementation and outcome evaluation stimulated by the models in each organisation. The learning emerging from this action research study suggests, in this respect, that the

development of programme theory can serve a requisite function in terms of the discussion and thinking around programme evaluation that was stimulated by its development.

Defining the role of developing programme theory to improve future programme evaluation practice as a requisite model is already considered in the literature on programme theory. In Chapter 2, I highlighted Patricia Rogers' citation of Weick, who said that even if the model is not entirely correct, that it can still provide a useful heuristic for purposeful action. This is because "once people begin to act … this helps them to discover what is occurring, what needs to be explained, and what should be done next" (Weick, 1995, pp.54–55). In Weick's words, having some form of model brings enough order to the world, which can then stimulate action (Weick, 1995). The learning emerging from this empirical action research study is consistent with, and supports, this thinking empirically through developing the understanding of *how* developing programme theory can stimulate evaluative thinking, even when the model of programme theory is either in a more draft form, or not entirely 'accurate'.

According to Phillips however, a model can be considered requisite only if no new insights emerge about the problem from the model (1984). Comparing the development of programme theory in this way, with the notion of a requisite model, is challenging because Phillips talks primarily about some sort of quantitative model, such as a cost-benefit analysis model. The findings presented in Chapter 5 suggest that, in both cases, the programme theory served a confirmatory function in that it presented information that participants stated they already knew. However, what the programme theory did was present familiar information in a *new* way and as such illuminated aspects of familiar information in ways which stimulated thinking. Moreover, it is difficult to say that if we had spent more time developing and validating the programme theory model, that new insights would not emerge and therefore the

programme theory would not be considered a requisite model by Phillips' definition.

Nevertheless, the learning emerging from the action research study suggests that the models of programme theory, in some capacity, were *sufficient,* at the time, to think more purposefully and constructively about programme implementation and outcome evaluation. In such cases, according to the concept of a requisite model, it is not that the programme theory necessarily captured the full reality of the programme, rather the programme theory model sufficiently reflects the shared social reality of those involved in developing and working with the model, including their own judgements on the reality which is being captured in the model (Phillips, 1984). In this way, the fact that the model reflected the shared social reality (through its confirmatory function and visual nature) was suffice to allow participating stakeholders to think about programme evaluation in more constructive and purposeful ways.

6.3.2 Development of programme theory as a transitional object

In both cases of action research, the ability of participants to think more constructively and purposefully about programme implementation and outcome evaluation was partly a result of its visual and tangible nature. In this way, the models of programme theory acted as physical artefacts that reflected the participants' viewpoints. This is similar to the concept of a transitional object, that emerged from the psychology field, when psychologists explored the benefits of learning through play, using objects that represented some part of reality e.g. a doll (Winnicott, 1989). In the case of modelling, a transitional object is said to reflect a 'microworld', or some social reality, that captures mental models that exist in the minds of those with whom you are working (De Geus, 1988). In this sense, a transitional object serves a confirmatory function in developing a common language and/or a shared understanding, for management teams, in the form of a visual object by making the programme theory explicit.

In both organisations, there was no formal validation of the extent to which the programme theory models created were an accurate reflection of the reality of programme implementation and outcomes i.e. we did not use them to test the programme theory. In Organisation 1, we used the perspectives of participants in the programme to create the model which might be more 'accurate' than say if it were developed solely by the management team, with whom I was working, as programme participants have lived experience of programme. However, I did not validate my maps with the young people, and so how accurately the maps reflect 'reality' is a function of my interpretation of what young people said, and how accurately they reflected on their experiences of the programme. Nevertheless, programme management were agreeable that the content of the programme theory confirmed their expectations of the programme theory. In Organisation 2, the maps were created by a team of programme stakeholders, some of whom have direct experience of the programme's implementation. In this case however, the map was purely developed from their own understanding, or mental models, of the programme, rather than direct lived experience.

Nevertheless, like in the case of requisite models, a transitional object need not accurately represent reality. Rather, in de Geus' words, "for the purposes of learning, it is not the reality that matters, but the team's model of reality, that will change, as members' understanding of their world improves" (1988). In other words, a *transitional object* is an object that is modified by participants and which therefore changes along with the alterations in participants understanding of the discussed problems. In such cases, it is the emerging shared understanding of the problem, reflected in the model, that stimulates action, rather than how accurately the model reflects some reality.

Moreover, the findings, emerging from Chapter 5, suggest that participants in both cases valued the visual nature of the programme theory maps through being able to see familiar information in a different way but also being able to actually see the pathways to intended changes in outcomes of the programmes they deliver. The importance of the visual nature of transitional models is also emphasised. The effective use of a transitional object is shown through the use of an object that reflects not just statements, but statements and their links, and as such this facilitates the creation of a collective view (Ackermann and Eden, 2011). In the cases presented in Chapter 5, the approach to mapping used focussed heavily on exploring the links between statements. In this way participating stakeholders were able to explore their understanding of how and why the programme was working through the visual links in the map rather than say with just narrative text (as would be the case in a case study, a commonly used method in formal (systematic) evaluation activities).

That being said, if those practising programme evaluation, in these two organisations, were to continue to utilise the programme theory to conduct formal evaluation activities, such as an outcome or implementation evaluation, it would mean that these mental models of programme stakeholders could be 'tested' against the 'real-world' implementation of a given programme, for example. It may be that in such cases, that programme stakeholders would discover that their mental models do not accurately reflect the reality of the programme. Nevertheless, for the purposes pursued by the organisations in this action research study, the development of models of programme theory were sufficient (requisite) as physical artefacts (transitional) to confirm a shared understanding and stimulate the thinking and discussion required to think more purposefully and constructively about programme implementation and outcome evaluation.

Overall, in using the concept of a transitional object to explain *how* developing programme theory can improve future programme evaluation practice in small TSOs, I argue that, in these cases, developing the model of programme theory gave participants a 'physical' object, or artefact, i.e. the visual map, to work with. The physical artefact depicted a common understanding (confirmed understanding) of the programme theory, which subsequently acted as a tool to think more purposefully and constructively about programme implementation and outcome evaluation activities.

6.3.3 Development of programme theory and model affordances

In both action research cases, programme theory models offered the opportunity to think more constructively and purposefully about programme implementation and outcome evaluation. The findings in Chapter 5 suggest that this was due to the confirmatory function and visual nature of the model of programme theory developed. In this sense, through the confirmatory function and visual nature of the models, we can think about *how* this *afforded* participants the opportunity to think about programme evaluation. This concept of 'affording opportunities' to take action relates to literature on models as *boundary objects* (Carlile, 2002, 2004; Franco, 2013) where models can offer model users the opportunity to navigate different boundaries they face relating to a specific problem situation. In the case of working with groups, these boundaries relate to the extent to which models generate shared meaning and interests through negotiation (Franco, 2013): this highlights the dual role of both the process of developing models, and model content, in ensuring model utility. There are three types of boundaries that models can help overcome, and in overcoming all three of these boundaries, a model becomes a boundary object. Carlile (2002) explains these boundaries in the following ways which I summarise below. I also consider what that would mean the case of models of programme theory.

1. Syntactic boundary – crossing the syntactic boundary is where there is a shared and stable language that enables communication. In the case

of developing programme theory, crossing the syntactic boundary would mean that there is already an explicit model of programme theory that is agreed upon by staff.

2. Semantic boundary – even when there is a shared and stable language, a semantic boundary occurs because interpretations of that language differ, i.e. there are differences in meaning, which can impact upon communication and collaborative working. Crossing the semantic boundary involves developing a shared understanding of the problem through making tacit knowledge explicit. In the case of developing programme theory, crossing the semantic boundary would involve making explicit the implicit, or tacit, programme theories (or mental models) and working with programme stakeholders develop a shared understanding of the programme theory.
3. Pragmatic boundary – when the interactions between people have consequences that mean people have something 'at stake' in negotiations, and thus it can be difficult for people to negotiate and change their mind. In the case of developing programme theory, it might be that the differences in meaning associated with the programme theory can have consequences for and on various stakeholders. Crossing the pragmatic boundary in this case would involve using the developing of programme theory to understand those consequences and negotiating ways to deal with them. An example of the use of programme theory in this was is in the case of overcoming designed blindness (Friedman, 2001).

Whilst I did not set out to explore the extent to which the development of a model of programme theory could act as a boundary object, this literature offers some interesting explanation of the opportunities for crossing these boundaries, that models of programme theory can offer. This notion of 'model affordances' can help further explain why the development of programme theory models in this case supported participants in thinking constructively and purposefully about programme evaluation. Franco (2013) identifies five model

affordances affecting the extent to which models can help users navigate different boundaries. These are listed below:

1. Tangibility – how a model makes its content visible to stimulate group discussion and negotiation
2. Associability – how a model can relate contents, based on shared characteristics, to identify where there are differences and dependencies in knowledge
3. Mutability – how a model can be modified, on the spot, to reflect changes based on discussion and negotiation
4. Traceability – how a model relates its content, with respect to time and structure of content, in order to explore content in more detail
5. Analysability – the transformation of model inputs to outputs

In both of the action research organisations, several of these affordances were illustrated relating to the development of programme theory, its confirmatory function, and visual nature. In both cases, participants observed that the information presented in the programme theory was largely unsurprising and confirmed how they thought the programme was working yet *afforded* those involved possibilities for action.

Examples of the affordances offered by the model of programme theory are numerous. The findings in Chapter 5 suggest that participants benefited from the visual nature of the maps both in illuminating familiar information in a new way (Organisation 1 and Organisation 2) and in visualising the pathways between programme delivery and multiple levels of outcomes (Organisation 2). As such this could suggest that the models of programme theory had high tangibility as well in that the visual afforded the participants the opportunity to discuss the contents of the model in terms of programme evaluation. Mutability, traceability, and analysability were also illustrated to a certain extent in the organisations. For example, in Organisation 2, we developed the programme theory in a workshop setting and as such were continually working

with, updating, and developing the map of programme theory. However, when it came to discussing programme evaluation, the map did not get updated much, as the already developed map seemed sufficient to work with. In both examples, the models did reflect some traceability and analysability in that with the SODA methodology, it was the structure of the map that we focussed on analysing. In Organisation 1, we discussed the analysis of maps in detail; however, it was not the participants who necessarily conducted the analysis. Rather, the analysis was presented back to the participants and used in the form of a facilitated discussion.

6.3.4 Role of facilitation

My role, in terms of facilitating the development and subsequent discussion of the programme theory models in each organisation, is important to consider. Whilst I was providing a resource to the organisations in terms of suggesting an tool we could take to develop programme theory (SODA methodology) and technical understanding of this tool, there are other important aspects to my role which I wish to consider, particularly in light of the preceding sections on the role of the model and modelling.

In this sense, my role as supporting evaluation could also be considered as a 'facilitated modeller' where facilitated modelling is the process by which formal models are jointly developed with a client group, face-to-face, with or without the assistance of a computer (Franco and Montibeller, 2010) in order to stimulate 'designed conversations'. Such conversations have the goal of generating increased understanding of a problematic situation and the development of commitments to consequential action (Franco, 2006). However, it is not the model alone that can facilitate such discussions. Rather, as in the cases presented in Chapter 5, my role was important to ensure that the models were developed, interpreted, and used in such a way that facilitated the discussion and subsequent suggestions for action about programme

implementation and outcome evaluation. In this sense I acted as a modeller and a facilitator.

In evaluation, this idea of a facilitated modeller places my role in each organisation as more of a change agent, i.e. an evaluator focussed on bringing about discussion, debate, and positive organisational change (Morabito, 2002), rather than as a technical evaluation expert. Whilst the literature on the development of programme theory certainly emphasises the role of stakeholder involvement, this notion of a facilitated modeller takes the role of stakeholder involvement beyond just having an input into the programme theory, rather than stakeholders are actively involved in both the development and subsequent use of the programme theory. This conceptualisation of the role of stakeholders sits better in cases where the actual systematic evaluation activities are undertaken by accidental evaluators. This stands in contrast to a situation where an external evaluator involves stakeholders in the development of programme theory, but goes on to carry out a formal evaluation separately from the stakeholders' involvement.

There are consequences to conceptualising my role in this research as a facilitative modeller in terms of the skills required. These skills include active listening, chart-writing (model building), managing group dynamics and power, and reaching closure (Franco and Montibeller, 2010). Many of these skills are associated with approaches to evaluation which aim to build capacity within and empower evaluation stakeholders. Such approaches include, but are not limited to, utilisation-focussed evaluation (Patton, 1997, 2008), empowerment evaluation (Fetterman, 1994; Fetterman and Wandersman, 2007) and democratic evaluation (Patton, 2002) where the goal of evaluation processes is to empower evaluation users and stakeholders in increasing their capacity to do and use evaluation. This is a narrative that exists in the literature on the use of programme theory in evaluation practice. For example, Patton himself takes a utilisation-focussed approach to using contribution analysis (Patton,

2012). However, this is certainly not a primary discourse in this literature, and perhaps with a more central position, such narratives would support a better understanding of *how* the development and subsequent use of programme theory could improve programme evaluation practices, particularly with respect to facilitate evaluative thinking.

6.3.5 Programme theory as a 'tool for evaluative thinking'

Overall, the findings, emerging from the action research component of this thesis, highlight the role that can be played by the development of programme theory in helping accidental evaluators to think more constructively and purposefully about programme evaluation. The findings discussed above are summarised in Table 9, which describes the main features of models, and modelling, and how they link to these findings. Table 9 can therefore serve as the beginnings of a conceptual framework to support future empirical studies seeking to explore the links between the development of programme theory, evaluative thinking, and formal (systematic) evaluation activities.

Model features	Aim	Link to chapter 5 findings
Requisite model	Sufficiency to solve a problem through shared view of social reality	Confirmatory function/enough information to work with
Transitional object	Physical artefact depicting a shared language	Visual nature & confirmatory function
Model affordances	Ability to do something with a model/opportunity for action	Conversations & model as a communication tool
Facilitated modelling	Designed conversation	Role of evaluator to facilitate thinking and change

Table 9 Summary of modelling literature links

6.3.6 Benefits of SODA mapping

At this point, it is important to reflect on the distinct benefit of using the SODA mapping approach in both organisations, as opposed to any other approach to developing programme theory. There are various approaches to developing programme theory (see table in Appendix 1), and as such it could be questioned why I did not choose to use one of those approaches instead. The reason for adopting the SODA approach was because it was a familiar tool for me, and one which I knew afforded flexibility in application, particularly in dealing with ill-defined problems. Flexibility in application was important as I was aware that the evaluation needs in each organisation would be different, and that at least using the same overall approach would afford some consistency.

There were three characteristics of the SODA mapping approach which were of benefit when working with the organisations. There were three key beneficial characteristics:

1. A structured process – by having a structured process to develop the programme theory, i.e. in terms of brainstorming, writing concepts, structuring concepts, and linking concepts, enabled different things in each organisation. In organisation 1, this process facilitated a structured process of map development and analysis which could then serve as a basis for discussion. This process in this sense could be transparently communicated to stakeholders, and thus more easily interpreted. In organisation 2, the structured process enabled us to develop the programmes theory in a way which was clear to stakeholders, and which facilitated developing the programme theory in a relatively short space of time.
2. 'Flexible' thinking – the use of SODA mapping allowed participants to think in a less linear and reductive way than might be implied by working with a linear box and arrow diagram. In this way, participants had the opportunity to explore specific issues in more detail, whilst retaining a sense of the bigger programme picture. Moreover, exploring the visual features of the maps, e.g. the ability to use multiple arrows in and out

of concepts, as well as explore loops, facilitated thinking about programmes as non-linear, and more reflective of the programme reality.

3. Use of language familiar to stakeholders and beneficiaries – working with concepts that make use of the stakeholders and/or beneficiaries' own language seemed to make the information within the maps more familiar and tangible. This stands in contrast to, for example, language that is too formalised or technical with respect to outcomes and programme implementation. Moreover, this use of familiar language within the visual maps illuminated issues that stakeholders were aware of but in ways which enabled them

6.4 A conceptual framework of the use of programme theory in programme evaluation practice in small TSOs

This chapter has presented a discussion on the findings of the empirical component of this thesis, with respect to both the extant knowledge in this area and drawing on additional literature to help explain the emergent learning from these two studies. In doing so, this thesis has developed a better understanding of the current use of programme theory in small TSOs, as well as how its use can improve future programme evaluation practice. I discuss that, whilst it is unsurprising that evaluation practitioners in small TSOs are both constrained by resources constrained and by the expectations of funders, that there is another important factor to consider that can influence the use of programme theory in formal (systematic) evaluation activities. This additional factor is that the use of programme theory is tacit in nature, in that it can be defined as 'knowledge-in-action' and, as such, may be difficult to articulate due to its embeddedness in practice. That being said, these three factors influence the current lack of use of programme theory in formal (systematic) evaluation activities, despite the learning about programme theory being of interest to both small TSOs and to funders. This leaves the question of what the value for small TSOs is, of making the programme theory explicit, and consequently

influencing the investment by both TSOs and funders of time and resources in developing it.

The learning from the action research study helps to conceptualise the answer to this question in addressing how the development of programme theory can improve future programme evaluation practice in small TSOs. The learning indicates that developing a model of programme theory can be in itself sufficient (as a requisite model) to stimulate evaluative thinking (i.e. more constructive and purposeful thinking around programme implementation and outcome evaluation) through its visual and confirmatory functions (as a transition object). In order to stimulate such constructive and purposeful thinking, models of programme theory can offer affordances to users in terms of their tangibility, mutability, associability, traceability and analysability, affordances which offer those involved opportunities to take action. Moreover, in supporting the development of programme theory in this way, those supporting the capacity to do evaluation in small TSOs, should re-evaluate their role in terms of how they can sufficiently support and facilitate the use of programme theory, through more effective knowledge transfer systems.

6.5 Chapter summary

This chapter has discussed the findings of the two studies that make up the empirical component of this thesis. The aim was to discuss the role of programme theory in small TSOs, both in terms of current practice and in terms of improving future practice. In drawing this discussion together at the end, I have presented a conceptual model of evaluation practice that aims to capture the role of programme theory in small TSOs from the practitioner perspective.

CHAPTER 7 CONCLUDING THE THESIS

7.1 Introduction

The concluding chapter of this thesis aims to consider the key implications of this research. First, I summarise the thesis, and the key thesis takeaways, before addressing the theoretical, methodological, and practical implications of the findings and discussion presented in the preceding three chapters. Specifically, I consider the implications for research on the use of programme theory in small TSOs; the implications for research on programme evaluation practice more generally; and the implications for programme evaluation practice in small TSOs. At the end of the chapter I consider the future research potential of the thesis, as well as the limitations, challenges and reflections on the research process.

7.1.1 Summary of thesis

Understanding social change in third sector activities is important due to the increasingly central role played by the third sector in tackling challenging social issues. In contexts such as the UK, the majority of third sector organisations are classed as small in size, and make up the majority, in number, of third sector organisations. Such organisations are usually based within local communities and contexts, and help to tackle some of the most challenging social issues facing society, such as youth disadvantage, social isolation, employment, health, amongst many others. As such, the role of programme evaluation is central to understanding social change efforts in communities, but also to ensure that those designing and delivering programmes can use evidence to inform their work. One relevant approach to programme evaluation, focussed on understanding how and why change occurs, is programme theory. As such, this thesis sought to explore the use of programme theory in small TSOs, where those conducting programme evaluation are likely to be ‘accidental evaluators’, with no formal training in programme evaluation.

This thesis presents research on evaluation practice in small TSOs with respect to the use of programme theory. In Chapter 1, I state that purposeful and systematic research on evaluation practice is important in order to close the gap the theory-practice gaps in programme evaluation, including with respect to the methods and approaches which are advocated in the evaluation literature. Such research aims to facilitate a better dialogue between how evaluation is actually conducted and how we believe evaluation should be done to understand better in which contexts and under which conditions different theories, methods, or approaches are useful and feasible to implement.

In addressing the use of programme theory in small TSOs currently, as well as *how* the development of programme theory can improve future programme evaluation practice in small TSOs, I adopt a multi-methodological framework that allows me to address the descriptive and prescriptive research questions respectively. Overall, I find that whilst the use of programme theory currently appears to be tacit in nature, rather than in any explicit form for use in formal (systematic) evaluation activities, that there are other practical constraints on its use. These practical constraints include the resources (expertise, time and money) to make better use of such tools, and the influence of funder expectations in shaping systematic evaluation activities. Moreover, I find that the development of programme theory can improve future programme evaluation, particularly in how it facilitates evaluative thinking. I conceptualise how the development of programme theory can facilitate evaluative thinking in terms of the role which the development of models, and modelling (of programme theory), can play in organisations (requisite, transitional, and affordances). Below I summarise the key thesis takeaways which inform the thesis implications for theory, research, and practice:

- The importance of programme evaluation, and the use of programme theory more specifically, is clear in the case of small TSOs, where social

change efforts are increasingly contracted out to such organisations because they are best placed to tackle them

- Different conceptualisations of evaluation practice influence what research on evaluation focuses on as well as the how research on evaluation is conducted
- The literature on the use of programme theory does not conceptualise evaluation practice in such a way that is consistent with how we understand programme evaluation practice in small TSOs
- Current use of programme theory in small TSOs can be defined as tacit in nature, relating specifically to the use of programme theory as 'knowledge-in-action'
- But there is an important role for funders in developing reporting and support mechanisms that support the use of programme theory in evaluation activities, through systems which tackle knowledge transfer issues
- However, within small TSOs, the development of programme theory can support future programme evaluation activities through its role as a 'tool for evaluative thinking'
- As a tool for evaluative thinking, programme theory can offer several benefits and opportunities through its confirmatory function and visual nature
- In supporting, the use of programme theory, the role of change agent (or facilitative modeller) is important to facilitate the use of such tools by accidental evaluators
- Taken together, defining the current use of programme theory as tacit in nature alongside the understanding of how developing programme theory can improve future programme evaluation practice, this thesis defines the beginnings of a conceptual model to support further empirical study in this area. Moreover, the methodological approach taken in this thesis can inform future pieces of research on evaluation through its ability to inform an effective dialogue between theory and

practice. This dialogue has further implications for supporting the practice of evaluation in small TSOs.

7.2 Implications for research on the use of programme theory in programme evaluation practice in small TSOs

This thesis has a number of theoretical implications that relate to improving our understanding of the use of programme theory in programme evaluation practice in small TSOs. In the first instance, the findings presented in this thesis, have better conceptualised the understanding of the current use of programme theory in small TSOs. Much of the previous literature identifies that there are practical constraints on the capacity of small TSOs to conduct programme evaluation, and that accountability to funders can be burdensome. This is consistent with the findings in this thesis. However, in adding to this understanding, the findings presented suggest that it is insufficient to simply discuss and explore what small TSOs cannot do with respect to programme evaluation. I find that there is some use of programme theory, which is tacit in nature, and used in unsystematic programme evaluation activities. Moreover, there is an appetite for the kind of information associated with programme theory to be used in evaluation in more explicit ways, which leaves the questions of what is the value of making knowledge which is tacit in nature, explicit; and how can the development and use of programme theory be better supported through the grantee-funder relationship?

The findings presented in this thesis also contribute to a better conceptualisation of *how* the development of programme theory can improve programme evaluation practice in small TSOs. By better linking the role of developing an explicit model of programme theory to the concepts of evaluative thinking (requisite, transitional, and affordances) and improved programme evaluation practices, this thesis presents the beginnings of a conceptual framework for better understanding the role of programme theory in such organisations as a 'tool for evaluative thinking'. Rather than simply a

method for conducting formal (systematic) evaluation activities, the findings presented in this thesis better conceptualise our understanding of the role that can be played by developing even a simple model of programme theory in small TSOs. This learning can be used to inform future studies on the use of programme theory in small TSOs. From an action research perspective, the academic community can utilise this framework to build up more empirical examples of developing programme theory, which are purposeful and systematic in nature. Doing so would enable the development of a larger portfolio of examples that can help to further refine the conceptual framework developed here and, consequently, better inform guidelines on the use of programme theory, particularly in the small TSO setting. Moreover, this conceptual framework could be used to support a better understanding of in which evaluation settings the development of programme theory is not considered useful, necessary, or feasible. For example, it may be the case that this conceptual framework is only useful in cases where those conducting evaluation are accidental evaluators and require more support to think through how to go about evaluating the programmes and services they offer. Moreover, the beneficial characteristics of SODA mapping, discussed in Chapter 6, should be noted by those wishing to research, develop, use, or support the use of programme theory in the future.

Moreover, the theoretical perspective adopted in this thesis (the practitioner-oriented perspective) stands in contrast to the dominant perspective in the literature on the use of programme theory in programme evaluation practice. The practitioner-oriented perspective adopted enables the findings to more sensitively reflect programme evaluation practice conducted within the small TSO context. This stands in contrast to the more methodological, implementation, or instrumental rationality, focus of the extant literature on the use of programme theory. In this sense, the findings presented capture evaluation practice in this setting, as well as contributing to the inclusion of more diverse perspectives in the wider evaluation field. Particularly in the case of the use of evaluation theory (in this case, programme theory), the literature

should, I argue, be more mindful of the diversity in programme evaluation practice, i.e. that not all evaluators are formally trained in evaluation, nor is programme evaluation practice merely theory enacted, that there are wider social and organisational contextual factors that influence evaluation practice.

Overall, in terms of an academic contribution, the learning generated in this thesis, supplements the extant literature, but accounts for a different perspective on programme evaluation practice. In doing so, I argue that the only way to actually change practice is to move beyond 'how-to' guides on the use of programme theory and the tracking of their methodological implementation, to an approach to research on evaluation which accounts for the diverse landscape of evaluation practice.

7.3 Implications for research on programme evaluation practice more generally

This thesis also has several more general methodological implications in terms of conducting research on programme evaluation. In particular, I suggest that research on evaluation practice should take a more systematic and purposeful methodological approach, echoing previous calls in the field. Like much of the evaluation literature, the literature on the use of programme theory is primarily based on individual reports of evaluations. Whilst such reports can be insightful and provide much detail, they are not necessarily purposeful or systematic (as given in the definition on research on evaluation presented in Chapter 1). As such, due to a lack of standard conceptual and methodological framework, it is often difficult to draw out learning about the use of specific evaluation theories from across studies. For example, in some of the reviews of the use of programme theory, it was unclear what the added value of the use of programme theory was, which was likely due, in part, to variability in reporting of its use. In the case of programme theory, it may be that, where the programme theory played an important part in the evaluation, it is not commented on in great detail in the report of the evaluation. Therefore,

systematic and purposeful inquiry, on the use of evaluation theory, can help us to understand the specific role of that evaluation theory. Of course, in the case of programme theory there are such studies, which mainly come in the form of systematic or structured reviews of the use of programme theory. However, such studies focus on methodological implementation of programme theory approaches, and omit the perspective of the practising evaluator.

Moreover, the multi-methodological framework adopted enables the empirical component of the thesis to address both descriptive and prescriptive elements of the research problem with empirical data. I do this through the simultaneous use of qualitative interviews (descriptive) and action research (prescriptive) research methods. In this sense, the thesis can effectively inform a dialogue between the descriptive and prescriptive aspects of research on programme evaluation practice. This particular systematic and purposeful approach has distinct value over individual case reports of evaluations, or the systematic review of evaluation examples, both of which can be largely descriptive, and-or often anecdotal, in nature. This means that this thesis has been able to account for both the current use of programme theory within the small TSO context, as well as understand how the use of programme theory could potentially improve that practice in the future. Improving this dialogue means that we can better inform strategies to support the use of programme theory, and evaluation theory more generally, in programme evaluation practice. Future research on evaluation studies should actively consider the dialogue between descriptive and prescriptive elements and how these can inform each other through the use of appropriate methodological frameworks.

More generally, another implication for future studies on the use of evaluation theory in programme evaluation practice, is the value of adopting a variety of perspectives, and in acknowledging the diversity of programme evaluation practice contexts. The literature on the use of programme theory in evaluation practice is predominantly from a technical perspective, placing the method and

its implementation as the focus of inquiry and analysis. Whilst this thesis just considers the case of programme theory, I imagine that if I chose to focus on other specific methods or approaches to conducting evaluation, similar to programme theory, I might find a similar trend. This study however, has adopted a practitioner-oriented perspective, and a methodology that places the practising evaluator, as the empirical focus. Both qualitative interviewing and the use of action research has allowed evaluation practitioners in small TSOs to both reflect-on and reflect-in their programme evaluation practice, adding value to insights generated about evaluation practice. As such, the learning generated in this thesis provides a better understanding of several factors, including: the contextual influences on the use of programme theory, the evaluation challenges the development of programme theory can help overcome, the ways in which it can help address those challenges, and the additional value it can have in thinking about programme evaluation more generally. By adopting a practitioner-oriented perspective, research on evaluation more generally can offer alternative perspectives on the use of evaluation theory through highlighting, for example, contextual influences on evaluation practice, or practitioners' implicit theories of evaluation practice.

Moreover, the practitioner focus adopted in this study means that the findings are likely more accessible, interesting, and relevant for those practising evaluation in small TSOs, or for those who support the practice of evaluation in similar organisations. From a more methodological perspective, the literature on the use of programme theory in practice could, for those conducting programme evaluation with limited evaluation expertise or knowledge of terminology, be intimidating, and subsequently restrict the use of such approaches. Using a practitioner perspective, this study serves to highlight the added value of developing programme theory from the perspective of those actually practising evaluation within small TSOs. With this, I hope that the findings of this thesis can more adequately address the gaps between theory and actual practice in the context of small TSOs.

7.4 Implications for programme evaluation practice in small TSOs

The findings emerging from the empirical component of this thesis give us a clearer sense of the needs of small TSOs in terms of the use of programme theory, as well as how to support the use of programme theory in this setting. In this study, I find that there is added value in terms of stimulating evaluative thinking through the process of developing programme theory. Making what was already known, but tacit in nature, explicit, enables organisations to think more constructively and purposefully about programme evaluation. This study highlights the value to such organisations of developing programme theory, in its role as a 'tool for evaluative thinking'. It may be that in such cases, this is an end in itself. It might be that simply making the programme theory explicit is enough to stimulate evaluative thinking with respect to planning formal (systematic) evaluation activity, and thus thinking more purposefully and constructively about programme implementation and outcome evaluation. Efforts to develop programme theory should also be mindful of the characteristics of the mapping approaches used and how to facilitate evaluative thinking through a structured processes, that allows for non-linear thinking, and which enables the use of terminology that is familiar to stakeholders.

Considering the role of programme theory, in the way described above, also has practical implications with respect to programme evaluation support, training, and capacity building efforts in small TSOs. In building capacity, it is important to ensure such efforts are proportionate to, and reflective of the needs, of programme evaluation in such organisations. There is an important role to be played for funders of, and those who support evaluation capacity within, small TSOs. Funders have an important role to play in developing more effective systems for eliciting the kinds of information which the use of programme theory can facilitate. Moreover, in supporting small TSOs, it is recommended that those supporting capacity-building serve a facilitative and

change-agent type role, where, like in the action research component of this thesis, the role of facilitating is an dynamic and active one, that involves skills such as listening, chart-writing (model building), managing group dynamics and power, and reaching closure, for example. This stands in contrast to traditional training formats, which come in the form of written or online guidelines, or in seminar format. Whilst it is acknowledged that providing this kind of support to all small TSOs is probably unrealistic, the findings suggest that, those who support evaluation capacity building in such organisations, should reconsider how the support evaluation capacity, and the skills required to do so at scale. There is an opportunity here to better align the links between specific methods, such as programme theory, and approaches to evaluation which embed the notions of facilitation, groups dynamics, and change-agent type roles (Patton, 1997, 2002, 2008; Fetterman, 1994; Fetterman and Wandersman, 2007)

In understanding better *how* the development of programme theory can stimulate evaluative thinking, it could be that those developing training materials, whether that be documents or workshops, can be more sensitive to the role of programme theory in this way. For example, it might be that rather than simply stating that developing programme theory can support evaluation activities through understanding how and why programmes work, that training materials can explain how such models facilitate evaluation activities, through the role of programme theory as a 'tool for evaluative thinking'. Moreover, it may be that interactive training sessions can better focus on the 'doing', rather than simply learning about an approach, the value of which was illustrated in the use of action research methods. As was demonstrated in Organisation 2, a half-day can be rather productive in at least pulling together a simple model of programme theory. There was value for participants in working in an interactive way to develop the programme theory. In cases such as Organisation 1, where more detailed analyses were required, it could be that such needs are better supported on a one-to-one basis.

A further practice-based implication of this thesis concerns the teaching of programme evaluation. The findings of this thesis suggest that the teaching of evaluation should strive to challenge the rhetoric around programme evaluation. Teaching should both focus on methodology but also place emphasis on the diversity of programme evaluation practice across different organisation types and in different settings. It follows that the evaluation literature and academic community can also be more mindful of ensuring a diversity of perspectives are included in research, particularly in literature that focuses on the use of specific methods, approaches, or tools to conduct evaluation, and the benefits or drawbacks their use has in specific settings or contexts. Only then will the academic community (both teaching and research) be able to more effectively inform a dialogue between descriptive and prescriptive theory, between theory and practice, and between and amongst practitioners.

7.5 Publication potential & dissemination

Whilst the empirical component of this thesis has not yet been published in an academic outlet, I certainly consider there to be value and potential in doing so, in terms of the theoretical, methodological, practical implications discussed in the preceding sections. Each of the empirical studies carried out generates its own insights and therefore could be published in its own right. Both studies present insights and value which would be of interest for both the evaluation and the third sector communities. Either of these academic outlets would seem plausible for publication. However, in terms of academic dissemination, I feel like this learning has more value, and potential impact, in an evaluation outlet as I think this research can challenge some of the limitations and dominant perspectives in the evaluation literature, particularly with respect to adopting a practitioner- oriented perspective, accounting for the diversity of evaluation practice, e.g. accidental evaluators and evaluation practice contexts, and in

more sensitively promoting the use of specific methods, approaches, or tools to elicit for those conducting evaluation.

Whilst the primary purpose of a PhD thesis is to contribute to 'knowledge', I believe that in the field of programme evaluation theory and practice come hand-in-hand. Therefore, I believe it can be a responsibility of academics, conducting research on evaluation, to share the insights and learning generated with the practice-based community, not all of whom will be engaged in the scholarly literature. Therefore, I perceive wider dissemination to be an important aspect of academic research on evaluation practice. Throughout the course of my PhD I have sought to engage with practitioners in this way. I have presented the insights to practitioner communities through practitioner focussed seminars, conferences, and written outlets[23]. I also shared, where appropriate, some of my knowledge with participants in the research. However, I think there is more integration between academic research and practice to be done, particularly in terms of integrating insights, such as those generated in this thesis, with programme evaluation guidelines, training materials, and evaluation support organisations. In particular, I think it is the task of those conducting research on evaluation to ensure that academic dissemination tools are accessible to the practitioners to whom they are most relevant.

7.6 Limitations, challenges, and reflections

First, there are a number of methodological limitations I wish to note. One of the primary limitations I experienced while conducting this study using an action research methodology concerns organisational time and resource demands (Morton, 1999). Certainly, this study would have benefited from more longitudinal follow-up with participating stakeholders, regarding understanding the added value of developing and using programme theory over the longer-

[23] https://www.evaluation.org.uk/app/uploads/2019/12/The-Evaluator-Autumn-2019.pdf

term; however, I have to appreciate that the time and resources of participating stakeholders are already stretched due to their busy roles within their respective organisations. I accept that the lack of a longer-term follow-up somewhat reflects the reality of the nature of the work, and programme evaluation practice, in the organisations with whom I was working. Nevertheless, by using an action research approach, I was able to gain access to insights which, otherwise, I would have been unable to attain using other research methods.

A related challenge in action research is maintaining a balance between the academic focus and quality of the research, and consultancy and benefits for the 'client' (Morton, 1999). In order to address this challenge, I attempted to the best of my ability to practice reciprocity (Robertson, 2000). Practising reciprocity refers to ensuring mutual benefit and negotiation of the meaning of the research, i.e. that the process of theory-building should be mutually beneficial (Robertson, 2000). In order to do this, I made the goals of my research clear from the outset, so that all participating stakeholders were aware of what I wanted to achieve from this study. I also ensured that I was clear about what could be achieved for the organisation, and what could not be achieved, in order to manage expectations of the process, keeping expectations realistic for both parties. Moreover, allowing participants to question and think about findings through asking stakeholders to review the write-up of this study through reports, and incorporating their feedback, further ensured the mutual benefit.

At the theoretical level, there were a number of challenges in terms of understanding and contributing to our understanding of programme evaluation practice. First, in Study 2, it is difficult to say to what extent it was the actual developing of programme theory that stimulated the discussions and thinking for participating stakeholders, or whether it was simply the opportunity to talk about evaluation. I try, in the findings presented in Chapter 5, to keep

reflections as closely focused on the added value of the process of developing, and the resulting content of, the programme theory. In this light, I think this study does provide some evidence that the process of developing programme theory has several distinct benefits. Second, this study presents only two examples of developing programme theory. Whilst there were common discussion points emerging from both organisations, it will be beneficial in future to continue to collect similar examples of the development of programme theory, in ways which are both purposeful and systematic, rather than simply reports of case studies of evaluations. Last, in Study 1, one could question how reflective of actual practice participants accounts of their experiences with evaluation actually were. By adopting a friendly and approachable manner, paired with a piloted question guide, I am confident that participants felt they could be honest and open about their experiences. In fact, many of the participants made comments on how valuable it was to talk through some of their experiences with programme evaluation.

The perspective adopted in this thesis acknowledges that practice is emergent and dynamic. The perspective adopted also acknowledges that the findings presented and discussed are very much a function of my 'transaction', or interaction, with the participants in, and contexts of, each organisation. Therefore, I make no claims to accessing some reality or 'knowledge' or truth about programme evaluation practice, or the use of programme theory. Rather, I prefer the learning generated in this thesis to be considered as contributing to our 'knowing' about this research problem, knowing that can be developed and adapted over time. Nevertheless, I consider that the insights generated can facilitate the conduct of more purposeful research that is systematic in nature, and which facilitates a more effective dialogue between descriptive and prescriptive aspects of programme evaluation practice.

Moreover, the insights generated in this thesis also have value for evaluation practice particularly to support the evaluation training and capacity building in

small TSOs which is proportionate to, and useful for, their evaluation needs. Future research should therefore build on the insights generated and approach taken here, particularly in terms of building portfolios of examples of the benefits to developing programme theory as well as being aware of evaluation practice on the ground in this setting, using a framework which enables learning to be generated from across examples whilst being sympathetic to the nature of evaluation practice. I also accept that it may the case that over time, as the contexts and needs of the sector change, frameworks to guide research on evaluation, such as that developed in this thesis, will also need to be adapted and updated to reflect the dynamic and emerging context of evaluation practice in the third sector.

7.7 Concluding comments

This thesis has presented a case of research on evaluation with respect to the use of programme theory in programme evaluation practice in small TSOs. The findings presented and discussed have helped to better conceptualise the use of programme theory within this setting. In addressing both descriptive and prescriptive elements of this research problem, I have been able to effectively inform a dialogue between theory and practice, which has theoretical, methodological, and practical implications.

In future, it is the job of those conducting research on evaluation to ensure that research on evaluation studies are purposeful and systematic in nature, and are representative of, and sensitive to, the diversity of programme evaluation practice, in what is a highly multidisciplinary field. Without these more diverse perspectives, research on evaluation, and the evaluation field more generally, risks being, in Christie's words, constrained by the prescriptive ideas on which the field is built, without adequate consideration of 'real-world' practice on the ground.

REFERENCES

Ackermann, F. and Eden, C., 2010. Strategic Options Development and Analysis. In: M. Reynolds and S. Holwell, eds., *Systems Approaches to Managing Change: A Practical Guide*, Springer. London, pp.135–190.

Ackermann, F. and Eden, C., 2011. Negotiation in Strategy Making Teams: Group Support Systems and the Process of Cognitive Change. *Group Decision and Negotiation*, 20(3), pp.293–314.

Alcock, P., 2010. A strategic unity: defining the third sector in the UK. *Voluntary Sector Review*, 1(1), pp.5–24.

Alkin, M.C., 2003. Evaluation theory and practice: Insights and new directions. *New Directions for Evaluation*, 2003(97), pp.81–90.

Argyris, C., 2003. Actionable Knowledge. In: H. Tsoukas and C. Knudsen, eds., *The Oxford Handbook of Organization Theory: Metatheoretical perspectives*. Oxford: Oxford University Press, pp.423–452.

Argyris, C. and Schön, D.., 1992. *Theory in practice: Increasing professional effectiveness*. Michigan: Jossey-Bass Inc.

Argyris, C. and Schön, D.., 1996. *Organisational Learning II*. Reading: Addison-Wesley Publishing Company.

Arvidson, M., 2009. *Impact and evaluation in the UK third sector: reviewing literature and exploring ideas*. *Third Sector Research Centre*, .

Astbury, B. and Leeuw, F.L., 2010. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation*, 31(3), pp.363–381.

Bach-Mortensen, A.M. and Montgomery, P., 2018. What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review. *Systematic Reviews*, 7(1), pp.1–15.

Bellis, M.A., Lowey, H., Leckenby, N., Hughes, K. and Harrison, D., 2013. Adverse childhood experiences: retrospective study to determine their impact on adult health behaviours and health outcomes in a UK population. *Journal of Public Health*, 36(1), pp.81–91.

Biesta, G.., 2009. How to Use Pragmatism Pragmatically? Suggestions for the Twenty-First Century. In: A.. Rud, J. Garrison and L. Stone, eds., *Dewey at 150: reflections for a new century*. West Lafayette: Purdue University Press.

Birckmayer, J.D. and Weiss, C.H., 2000. Theory-based evaluation in practice: What do we learn? *Evaluation Review*, 24(4), pp.407–431.

Blamey, A. and Mackenzie, M., 2007. Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges? *Evaluation*, 13(4), pp.439–455.

Bradbury, H., 2015. Introduction: how to situate and define action research. In: H. Bradbury, ed., *The SAGE Handbook of Action Research*, 3rd ed. London: SAGE Publications Ltd.

Breuer, E., Lee, L., De Silva, M. and Lund, C., 2016. Using theory of change to design and evaluate public health interventions: A systematic review. *Implementation Science*, 11(63).

Brewer, J. and Hunter, A., 2006. *Foundations of Multimethod Research: Synthesising Styles*. Thousand Oaks, California: SAGE Publications, Inc.

Brousselle, A. and Champagne, F., 2011. Program theory evaluation: Logic analysis. *Evaluation and Program Planning*, 34(1), pp.69–78.

Bryman, A., 2016. *Social Research Methods*. 5th ed. Oxford: Oxford University Press.

Buckley, J., Archibald, T., Hargraves, M. and Trochim, W.M., 2015. Defining and Teaching Evaluative Thinking: Insights From Research on Critical Thinking. *American Journal of Evaluation*, 36(3), pp.375–388.

Carlile, P.R., 2002. A pragmatic view of knowledge and boundaries: Boundary objects in new product development. *Organization Science*, 13(4), pp.442-455+456.

Carlile, P.R., 2004. Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science*, 15(5), pp.555–568.

Carman, J.G., 2011. Understanding Evaluation in Nonprofit Organizations. *Public Performance & Management Review*, 34(3), pp.350–377.

Carman, J.G. and Fredericks, K.A., 2008. Nonprofits and evaluation: Empirical evidence from the field. In: *Nonprofits and evaluation. New Directions for Evaluation*. pp.51–71.

Carman, J.G. and Fredericks, K.A., 2010. Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation*, 31(1), pp.84–104.

Center of Disease Control and Prevention, 2019. *Types of Evaluation*. [online] Center of Disease Control and Prevention. Available at: <https://www.cdc.gov/std/Program/pupestd/Types of Evaluation.pdf> [Accessed 10 Sep. 2019].

Chelimsky, E., 2006. The purposes of evaluation in a democratic society. In: I.. Shaw, J.C. Greene and M.. Mark, eds., *The SAGE Handbook of Evaluation*, 1st ed. London: SAGE Publications Ltd, pp.33–55.

Chen, H. and Rossi, P.H., 1980. The Multi-Goal , Theory-Driven Approach to Evaluation : A Model Linking Basic and Applied Social Science. *Social Forces*, 59(1), pp.106–122.

Chen, H. and Rossi, P.H., 1983. Evaluating with sense: The Theory-Driven

Approach. *Evaluation Review*, 7(3), pp.283–302.

Chen, H.T. and Rossi, P.H., 1987. The Theory Driven Approach To Validity. *Evaluation and Program Planning*, 10, pp.95–103.

Chouinard, J.A., Boyce, A.S., Hicks, J., Jones, J., Long, J., Pitts, R. and Stockdale, M., 2017. Navigating Theory and Practice Through Evaluation Fieldwork. *American Journal of Evaluation*, 38(4), pp.493–506.

Christie, C., 2003a. Editor's Notes. *New Directions for Evaluation*, Spring 200(97), pp.1–6.

Christie, C., 2003b. Understanding evaluation theory and its role in guiding practice: Formal, folk, and otherwise. *New Directions for Evaluation*, Spring 200(97), pp.91–93.

Christie, C., 2003c. What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation*, 2003(97), pp.7–36.

Cooksy, L.J., Gill, P. and Kelly, P.A., 2001. The program logic model as an integrative framework for a multimethod evaluation. *Evaluation and Program Planning*, 24, pp.119–128.

Coryn, C.L.S., Ozeki, S., Wilson, L.N., Greenman, G.D., Schröter, D.C., Hobson, K.A., Azzam, T. and Vo, A.T., 2016. Does Research on Evaluation Matter? Findings From a Survey of American Evaluation Association Members and Prominent Evaluation Theorists and Scholars. *American Journal of Evaluation*, 37(2), pp.159–173.

Coryn, C.L.S., Wilson, L.N., Westine, C.D., Hobson, K.A., Ozeki, S., Fiekowsky, E.L., Greenman, G.D. and Schröter, D.C., 2017. A Decade of Research on Evaluation: A Systematic Review of Research on Evaluation Published Between 2005 and 2014. *American Journal of Evaluation*, 38(3), pp.329–347.

Coryn, C.L.S.S., Noakes, L.A., Westine, C.D., Schröter, D.C. and Schro, D.C., 2011. A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. *American Journal of Evaluation*, 32(2), pp.199–226.

Costa, E., Ramus, T. and Andreaus, M., 2011. Accountability as a Managerial Tool in Non-Profit Organizations: Evidence from Italian CSVs. *Voluntas*, 22(3), pp.470–493.

Cousins, J.B., Goh, S.., Elliott, C.. and Bourgeois, I., 2014. Framing the capacity to do and use evaluation. In: *New Directions for Evaluation*. pp.7–23.

Cousins, J.B., Goh, S.C., Clark, S. and Lee, L.E., 2004. Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *The Canadian Journal of Program Evaluation*, 19(2), pp.99–141.

Craig, P. and Campbell, M., 2015. *Evaluability Assessment : a systematic approach to deciding whether and how to evaluate programmes and policies*.

Creswell, J.. and Creswell, J.D., 2018. *Research Design: Qualitative, quantitative, and mixed methods approaches*. 5th ed. Thousand Oaks: SAGE Publications, Inc.

Dahler-Larsen, P., 2012. *The Evaluation Society*. 1st ed. Stanford, California: Stanford University Press.

Dahler-Larsen, P., 2014. Dynamite, Medicine, Peace, and the Nobel Art of Evaluation. *American Journal of Evaluation*, 35(3), pp.377–386.

Dahler-Larsen, P., 2016. The Changing Role of Evaluation in a Changing Society. In: R. Stockmann and W. Meyer, eds., *The Future of Evaluation Global Trends, New Challenges, Shared Perspectives*. Basingstoke: Palgrave Macmillan, pp.24–35.

Dalkin, S., Lhussier, M., Williams, L., Burton, C.R. and Rycroft-Malone, J., 2018. Exploring the use of Soft Systems Methodology with realist approaches: A novel way to map programme complexity and develop and refine programme theory. *Evaluation*, 24(1), pp.84–97.

Dalkin, S.M., Greenhalgh, J., Jones, D., Cunningham, B. and Lhussier, M., 2015. What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, 10(1), pp.1–7.

Davies, R., 2018. Representing theories of change: technical challenges with evaluation consequences. *Journal of Development Effectiveness*, 10(4), pp.438–461.

Deane, K.L. and Harré, N., 2014. Program theory-driven evaluation science in a youth development context. *Evaluation and Program Planning*, 45, pp.61–70.

Denzin, N.K., 2012. Triangulation 2.0*. *Journal of Mixed Methods Research*, 6(2), pp.80–88.

Dhanani, A. and Connolly, C., 2012. Discharging not-for-profit accountability: UK charities and public discourse. *Accounting, Auditing and Accountability Journal*, 25(7), pp.1140–1169.

Donaldson, S., 2007. *Program Theory-Driven Evaluation Science: Strategies and Applications*. New York: Lawrence Erlbaum Associates.

Ebrahim, A., 2003. Making sense of accountability: Conceptual perspectives for northern and southern nonprofits. *Nonprofit Management and Leadership*, 14(2), pp.191–212.

Eden, C., 1988. Cognitive Mapping. *European Journal of Operational Research*, 36, pp.1–13.

Eden, C., 1992. On the nature of cognitive maps. *Journal of Management Studies*, 29(3).

Eden, C., 2004. Analyzing cognitive maps to help structure issues or problems. *European Journal of Operational Research*, 159, pp.673–686.

Eden, C. and Ackermann, F., 1998. Detecting emergent strategizing: Working with strategy making teams - The oval mapping technique. *Making Strategy: The Journey of Strategic Management*, pp.303–320.

Eden, C. and Ackermann, F., 2001. SODA - The Principles. In: J. Rosenhead and J. Mingers, eds., *Rational Analysis for a Problematic World: Problem Structuring Methods for Complexity, Unertainty, and Conflict*, 2nd ed. Chichester: John Wiley & Sons, Ltd.

Eden, C., Ackermann, F. and Cropper, S., 1992. The Analysis of Cause Maps. *Journal of Management Studies*, 29(3), pp.309–324.

Eden, C. and Huxham, C., 1996. Action Research for Management Research. *British Journal of Management*, 7, pp.75–86.

Eden, C. and Huxham, C., 2002. Action Research. In: D. Partington, ed., *Essential Skills for Management Research*. London: SAGE Publications Ltd, pp.254–272.

Ellis, J. and Gregory, T., 2009. *Accountability and learning: Developing monitoring and evaluation in the third sector*. *Charities Evaluation Services*. London.

Evaluation Support Scotland, 1998. *Evaluation Support guide 1 . 2 Developing a Logic Model*. [online] Available at: <http://www.evaluationsupportscotland.org.uk/resources/127/> [Accessed 11 Dec. 2019].

Fetterman, D.., 1994. Empowerment Evaluation. *Evaluation Practice*, 15(1), pp.1–15.

Fetterman, D. and Wandersman, A., 2007. Empowerment evaluation: Yesterday, today, and tomorrow. *American Journal of Evaluation*, 28(2), pp.179–198.

Fine, A.H., Thayer, C.E. and Coghlan, A., 2000. Program evaluation practice in the nonprofit sector. *Nonprofit Management & Leadership*, 10(3), pp.331–339.

Franco, L.A., 2006. Forms of conversation and problem structuring methods: A conceptual development. *Journal of the Operational Research Society*, 57(7), pp.813–821.

Franco, L.A., 2013. Rethinking soft or interventions: Models as boundary objects. *European Journal of Operational Research*, 231(3), pp.720–733.

Franco, L.A. and Montibeller, G., 2010. Facilitated modelling in operational research. *European Journal of Operational Research*, 205(3), pp.489–500.

Frechtling, J.A., 2007. *Logic Modelling Methods in Program Evaluation*. San Francisco: Jossey-Bass.

Friedman, V.J., 2001. Designed Blindness: An action science perspective on Programe Theory Evaluation. *American Journal of Evaluation*, 22(2), pp.161–

181.

Funnell, S.. and Rogers, P.., 2011. *Purposeful Program Theory: Effective use of theories of change and logic models*. San Francisco: Jossey-Bass.

Funnell, S.C., 2000. Developing and using a program theory matrix for program evaluation and performance monitoring. *New Directions for Evaluation*, 2000(87), pp.91–101.

Galport, M. and Galport, N., 2015. Methodological trends in research on evaluation. *New Directions for Evaluation*, 148(Winter 2015), pp.17–29.

De Geus, A.P., 1988. *Planning as Learning.* [online] Harvard Business Review. Available at: <https://hbr.org/1988/03/planning-as-learning> [Accessed 24 Sep. 2019].

Greenhalgh, T., Thorne, S. and Malterud, K., 2018. Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation*, 48(6), pp.1–6.

Gugiu, P.C. and Rodríguez-Campos, L., 2007. Semi-structured interview protocol for constructing logic models. *Evaluation and Program Planning*, 30(4), pp.339–350.

Hacsi, T.A., 2000. Program Theory to Replicate Successful Programs. *New Directions for Evaluation*, (87), pp.71–78.

Hardwick, R., 2018. *Research use and knowledge mobilisation in third sector organisations involved in health care provision*. *Doctor of Philosophy in Medical Studies*. University of Exeter.

Hardwick, R., Anderson, R. and Cooper, C., 2015. How do third sector organisations use research and other knowledge? A systematic scoping review. *Implementation Science*, 10(1), pp.1–12.

Harlock, J., 2013. *Impact measurement practice in the UK third sector : a review of emerging evidence*. *Third Sector Research Centre*, .

Hatry, H., Van Houten, T., Plantz, M.C. and Greenway, M.T., 1996. Measuring program outcomes: A practical approach. *United Way of America: Measuring Program Outcomes*, pp.129–146.

Heckman, J.J., 2006. Skill formation and the economics of investing in disadvantaged children. *Science*, 312(June), pp.1900–1902.

Heckman, J.J., 2008. Schools, Skills, and Synapses. *Economic Inquiry*, 46(3), pp.289–324.

Heckman, J.J., 2012. The Developmental Origins of Health. *Health Economics*, 21(1), pp.24–29.

Hernandez, M., 2000. Using Logic-Models and program Theory to Build Outcome Accountability. *Education & Treatment of Children (ETC)*, 23(1), p.24.

Herr, K. and Anderson, G., 2015. *The Action Resreach Dissertation: A guide for students and faculty*. 2nd ed. Thousand Oaks, California: SAGE Publications, Inc.

Hesse-Biber, S., 2015. Introduction: navigating a turbulent research landscape: working the boundaries, tensions, diversity, and contradictions of multimethod and mixed methods inquiry. In: S. Hesse-Biber and B.R. Johnson, eds., *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. New York: Oxford University Press, pp.xxxiii–liii.

Hills, M., Carroll, S. and Desjardins, S., 2010. Asset Based Interventions: Evaluating and Synthesising Evidence of Effectiveness of the Assets Based Approach to Health Promotion. In: A. Morgan, M. Davies and E. Ziglio, eds., *Health Assets in a Global Context*. New York: Springer Science + Business Media, Inc, pp.77–98.

HM Treasury, 2018. *The Green Book: Central Government Guidance on Appraisal and Evaluation*. *The Green Book*.

Holma, K. and Kontinen, T., 2011. Realistic evaluation as an avenue to learning for development NGOs. *Evaluation*, 17(2), pp.181–192.

Huebner, T.A., 2000. Theory-based evaluation: Gaining a shared understanding between school staff and evaluators. *New Directions for Evaluation*, 2000(87), pp.79–89.

Huxham, C. and Vangen, S., 2003. Researching Organizational Practice Through Action Research: Case Studies and Design Choices. *Organizational Research Methods*, 6(3), pp.383–403.

Inspiring Impact, 2013. *Are you leading for impact? Five questions for voluntary sector leaders*.

Johnson, K., Greenseid, L.O., Toal, S.A., King, J.A., Lawrenz, F. and Volkov, B., 2009. Research on Evaluation Use: A Review of the Empirical Literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), pp.377–410.

Julian, D.A., 1997. The utilization of the logic model as a system level planning and evaluation device. *Evaluation and Program Planning*, 20(3), pp.251–257.

Julian, D.A., Jones, A. and Deyo, D., 1995. Open systems evaluation and the logic model: Program planning and evaluation tools. *Evaluation and Program Planning*, 18(4), pp.333–341.

Keller, A. and Bauerle, J.A., 2009. Using a logic model to relate the strategic to the tactical in program planning and evaluation: An illustration based on social norms interventions. *American Journal of Health Promotion*, 24(2), pp.89–92.

King, J.A., 2003. The challenge of studying evalutation theory. *New Directions for Evaluation*, 2003(97), pp.57–68.

Labin, S.N., Duffy, J.L., Meyers, D.C., Wandersman, A. and Lesesne, C.A., 2012. A Research Synthesis of the Evaluation Capacity Building Literature.

American Journal of Evaluation, 33(3), pp.307–338.

Lacouture, A., Breton, E., Guichard, A. and Ridde, V., 2015. The concept of mechanism from a realist approach: A scoping review to facilitate its operationalization in public health program evaluation. *Implementation Science*, 10(1), pp.1–10.

Lang, I.A. and Hardwick, R., 2016. Knowledge, knowledge mobilisation and the third sector. *Voluntary Sector Review*, 7(2), pp.209–218.

Langbein, L., 2012. *Public Program Evaluation*. 2nd ed. Armonk, NYC: M.E. Sharpe, Inc.

Layard, R., Clark, A.., Cornaglia, F., Powdthavee, N. and Vernoit, J., 2014. What predicts a successful life? A life-course model of well-being. *The Economic Journal*, 124, pp.720–738.

Leviton, L.C., 2015. Evaluation Practice and Theory: Up and Down the Ladder of Abstraction. *American Journal of Evaluation*, 36(2), pp.238–242.

Leviton, L.C., Khan, L.K., Rog, D., Dawkins, N. and Cotton, D., 2010. Evaluability Assessment to Improve Public Health Policies, Programs, and Practices. *Annual Review of Public Health*, 31(1), pp.213–233.

Lipsey, M.W. and Pollard, J.A., 1989. Driving Toward Theory in Program Evaluation More Models To Choose From. *Evaluation and program planning*, 12, pp.317–328.

Macmillan, R., 2010. The third sector delivering public services: an evidence review. Third Sector Research Centre Working Paper. Working Pa(July).

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T. and Kegels, G., 2012. Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18(2), pp.192–212.

Mark, M.M., Henry, G.T. and Julnes, G., 1999. Toward an Integrative Framework for Evaluation Practice. *American Journal of Evaluation*, 20(2), pp.177–198.

Mark, M.M., Henry, G.T. and Julnes, G., 2000. *Evaluation: An integrated framework for understanding, guiding and improving policies and programs*. San Francisco: Jossey-Bass Inc.

Markantoni, M., Steiner, A.A. and Meador, J.E., 2019. Can community interventions change resilience? Fostering perceptions of individual and community resilience in rural places. *Community Development*, 50(2), pp.238–255.

Mayne, J., 1999. *Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly*. *Discussion Paper*, .

McAdam, R., Mason, B. and McCrory, J., 2007. Exploring the dichotomies within the tacit knowledge literature: Towards a process of tacit knowing in organizations. *Journal of Knowledge Management*, 11(2), pp.43–59.

McLaughlin, J.A. and Jordan, G.B., 1999. Logic Models: A Tool for Telling Your Program's Performance Story. *Evaluation and Program Planning*, 22(1), pp.1–14.

Miles, M.. and Huberman, A.M., 1994. *Qualitative Data Analysis*. 2nd ed. Thousand Oaks: SAGE Publications, Inc.

Millar, A., Simeone, R.S. and Carnevale, J.T., 2001. Logic models: A systems tool for performance management. *Evaluation and Program Planning*, 24(1), pp.73–81.

Miller, R.L., 2010. Developing Standards for Empirical Examinations of Evaluation Theory. *American Journal of Evaluation*, 31(3), pp.390–399.

Mills, T., Lawton, R. and Sheard, L., 2019. Advancing complexity science in healthcare research: the logic of logic models. *BMC Medical Research Methodology*, 19(1), p.55.

Mintzberg, H., 1973. *The Nature of Managerial Work*. New York: Harper Collins Publishers, Inc.

Mintzberg, H., 2009. *Managing*. Harlow: Pearson Education Limited.

Morabito, S.M., 2002. Evaluator roles and strategies for expanding evaluation process influence. *American Journal of Evaluation*, 23(3), pp.321–330.

Moran-Ellis, J., Alexander, V.D., Cronin, A., Dickinson, M., Fielding, J., Sleney, J. and Thomas, H., 2006. Triangulation and integration: Processes, claims and implications. *Qualitative Research*, 6(1), pp.45–59.

Morgan, D.L., 2007. Paradigms Lost and Pragmatism Regained: methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), pp.48–76.

Morgan, D.L., 2014. Pragmatism as a Paradigm for Social Research. *Qualitative Inquiry*, 20(8), pp.1045–1053.

Morton, A., 1999. Ethics In Action Research. *Systemic Practice and Action Research*, 12(2), pp.219–222.

Olsson, C.A., McGee, R., Nada-Raja, S. and Williams, S.M., 2013. A 32-Year Longitudinal Study of Child and Adolescent Pathways to Well-Being in Adulthood. *Journal of Happiness Studies*, 14(3), pp.1069–1083.

Padgett, D.., 2008. *Qualitative Methods in Social Work Research*. 2nd ed. Thousand Oaks, California: SAGE Publications, Inc.

Palumbo, D.J. and Oliverio, A., 1989. Implementation theory and the theory-driven approach to validity. *Evaluation and Program Planning*, 12(4), pp.337–344.

Patton, M.., 1990. *Qualitative Evaluation and Research Methods*. 2nd ed. California: SAGE Publications, Inc.

Patton, M.., 1997. *Utilization-focused evaluation*. 3rd ed. Thousand Oaks:

SAGE Publications, Inc.

Patton, M.., 2008. *Utilization-Focused Evaluation*. 4th ed. Thousand Oaks: SAGE Publications, Inc.

Patton, M.Q., 1998. Discovering Process Use. *Evaluation*, 4(2), pp.225–233.

Patton, M.Q., 2002. A Vision of Evaluation that Strengthens Democracy. *Evaluation*, 8(1), pp.125–139.

Patton, M.Q., 2011. *Developmental Evaluation: applying complexity concepts to enhance innovation and use*. 1st ed. New York: The Guilford Press.

Patton, M.Q., 2012. A utilization-focused approach to contribution analysis. *Evaluation*, 18(3), pp.364–377.

Patton, M.Q., 2018. A Historical Perspective on the Evolution of Evaluative Thinking. *New Directions for Evaluation*, 2018(158), pp.11–28.

Pawson, R. and Tilley, N., 1997. *Realistic Evaluation*. London: SAGE Publications Ltd.

Peyton, D.J. and Scicchitano, M., 2017. Devil is in the details: Using logic models to investigate program process. *Evaluation and Program Planning*, 65(January), pp.156–162.

Phillips, L.D., 1984. A theory of requisite decision models. *Acta Psychologica*, 56(1–3), pp.29–48.

Polanyi, M., 1962. Tacit Knowing: Its bearing on some problems of philosophy. *Reviews of Modern Physics*, 34(4), pp.601–616.

Polanyi, M., 1966. The Logic of Tacit Inference. *The Journal of the Royal Institute of Philosophy*, 41(155), pp.1–18.

Power, M., 1997. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.

Preskill, H., 2014. Now for the Hard Stuff: Next Steps in ECB Research and Practice. *American Journal of Evaluation*, 35(1), pp.116–119.

Preskill, H. and Boyle, S., 2008. A Multidisciplinary Model of Evaluation Capacity Building. *American Journal of Evaluation*, 29(4), pp.443–459.

Preskill, H. and Torres, R.T., 1999a. Building capacity for organizational learning through evaluative inquiry. *Evaluation*, 5(1), pp.42–61.

Preskill, H.S. and Russ-Eft, D., 2016. *Building Evaluation Capacity*. 2nd ed. Thousand Oaks: SAGE Publications, Inc.

Preskill, H.S. and Torres, R.T., 1999b. *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, California: Sage Publications, Inc.

Pyrko, I., Dörfler, V. and Eden, C., 2017. Thinking together: What makes Communities of Practice work? *Human Relations*, 70(4), pp.389–409.

Reed, J.H. and Jordan, G., 2007. Using systems theory and logic models to define integrated outcomes and performance measures in multi-program settings. *Research Evaluation*, 16(3), pp.169–181.

Ricciuti, E. and Calo, F., 2016. *The state-of-the-art of Italian foundations' frameworks and methods to measure impact*. *Working paper n. 15*, .

Robertson, J., 2000. The three Rs of action research methodology: Reciprocity, reflexivity and reflection-on-reality. *Educational Action Research*, 8(2), pp.307–326.

Rog, D.J., 2015. Infusing Theory Into Practice, Practice Into Theory: Small Wins and Big Gains for Evaluation. *American Journal of Evaluation*, 36(2), pp.223–238.

Rogers, P.., Petrosino, A., Huebner, T.. and Hacsi, T.., 2000. Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation*, 87, pp.5–13.

Rogers, P. and McDonald, B., 1999. Three Achilles Heels of Program Theory Evaluation. In: *Australasian Evaluation Society conference*.

Rogers, P.J., 2008. Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1), pp.29–48.

Rosas, S.R., 2005. Concept mapping as a technique for program theory development: An illustration using family support programs. *American Journal of Evaluation*, 26(3), pp.389–401.

Rosenhead, J., 1996. What's the Problem ? An Introduction to Problem Structuring Methods. *Interfaces*, 26(6), pp.117–131.

Roy, M.J., Donaldson, C., Baker, R. and Kerr, S., 2014. The potential of social enterprise to enhance health and well-being: A model and systematic review. *Social Science and Medicine*, 123, pp.182–193.

Salamon, L.M. and Sokolowski, S.W., 2016. Beyond Nonprofits: Re-conceptualizing the Third Sector. *Voluntas*, 27(4), pp.1515–1545.

Saunders, M., Lewis, P. and Thornhill, A., 2016. *Research Methods for Business Students*. 7th ed. Harlow: Pearson Education Limited.

Scheirer, M.A. and Newcomer, K., 2001. Opportunities for program evaluators to facilitate performance-based management. *Evaluation and Program Planning*, 24(1), pp.63–71.

Schön, D.., 1991. *The Reflective Practitioner: How professionals think in action*. London: Ashgate Publishing Ltd.

Schwandt, T.A., 1989. Recapturing Moral Discourse in Evaluation. *Educational Researcher*, 18(8), pp.11–17.

Schwandt, T.A., 2003. `Back to the Rough Ground!' Beyond Theory to Practice in Evaluation. *Evaluation*, 9(3), pp.353–364.

Schwandt, T.A., 2005. The centrality of practice to evaluation. *American Journal of Evaluation*, 26(1), pp.95–105.

Schwandt, T.A., 2014. On the Mutually Informing Relationship Between Practice and Theory in Evaluation. *American Journal of Evaluation*, 35(2), pp.231–236.

Schwandt, T.A., 2018. Evaluative thinking as a collaborative social practice: the case of boundary judgement making. *New Directions for Evaluation*, 158, pp.125–137.

Scriven, M., 1991. *Evaluation thesaurus*. 4th ed. Newbury Park, California: SAGE Publications, Inc.

Shadish, W.R.J., Cook, T.. and Leviton, L.., 1991. *Foundations of Program Evaluation*. California: SAGE Publications, Inc.

Simpson, B., 2018. Pragmatism: A philosophy of practice. In: C. Cassell, A.L. Cunliffe and G. Grandy, eds., *SAGE Handbook of Qualitative Business and Management*. London: SAGE Publications Ltd, pp.54–68.

Smith, N.L., 1993. Improving Evaluation Theory Through The Empirical Study of Evaluation Practice. *American Journal of Evaluation*, 14(3), pp.237–242.

Soteri-Proctor, A., 2010. Making use of qualitative tools: Towards a fuller understanding of the voluntary sector's engagement with public service delivery. *International Journal of Social Research Methodology*, 13(5), pp.411–424.

Stame, N., 2004. Theory-Based Evaluation and Types of Complexity. *Evaluation*, 10(1), pp.58–76.

Steiner, A. and Teasdale, S., 2018. Unlocking the potential of rural social enterprise. *Journal of Rural Studies*, 70, pp.144–154.

Stringer, E., 2007. *Action Research*. 3rd ed. Thousand Oaks, California: SAGE Publications, Inc.

Sullivan, H., Barnes, M. and Matka, E., 2002. Building Collaborative Capacity through ' Theories of Change '. *Evaluation*, 8(2), pp.205–226.

Sullivan, H. and Stewart, M., 2006. Who Owns the Theory of Change? *Evaluation*, 12(2), pp.179–199.

Taylor-Powell, E., Jones, L. and Henert, E., 2003. *Enhancing Program Performance with Logic Models*.

Tenbensel, T.., Dwyer, J.. and Lavoie, J.., 2014. How not to kill the golden goose: Reconceptualizing accountability environments of third sector organisations. *Public Management Review*, 16(7), pp.925–944.

Thayer, C.E. and Fine, A.H., 2001. Evaluation and outcome measurement in the non-profit sector : stakeholder participation. *Evaluation and Program Planning*, 24, pp.103–108.

Trevisan, M.S., 2007. Evaluability assessment From 1986 to 2006. *American Journal of Evaluation*, 28(3), pp.290–303.

Valentine, K., 2016. Complex Needs and Wicked Problems: How Social Disadvantage Became Multiple. *Social Policy and Society*, 15(2), pp.237–249.

Vallin, L., Philippoff, J., Pierce, S. and Brandon, P., 2015. Research-on-evaluation articles published in the American Journal of Evaluation, 1998-2014. *New Directions for Evaluation*, (148), pp.7–15.

Van de Ven, A.H. and Johnson, P.E., 2006. Knowledge for Theory and Practice. *Academy of Management Review*, 31(4), pp.802–821.

Vo, A. and Christie, C., 2015. Advancing research on evaluation through the study of context. *New Directions for Evaluation*, (148), pp.43–55.

Vo, A.T., Schreiber, J.S. and Martin, A., 2018. Toward a Conceptual Understanding of Evaluative Thinking. *New Directions for Evaluation*, 2018(158), pp.29–47.

Vogel, I., 2012. *Review of the use of 'Theory of Change' in international development*.

Walsemann, K.M., Geronimus, A.T. and Gee, G.C., 2008. Accumulating Disadvantage Over the Life Course: Evidence From a Longitudinal Study Investigating the Relationship Between Educational Advantage in Youth and Health in Middle Age. *Research on Aging*, 30(2), pp.169–199.

Weick, K.E., 1995. *Sensemaking in Organizations*. Thousand Oaks, California: SAGE Publications, Inc.

Weiss, C.H., 1995. Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families. In: J.I. Connell, A.C. Kubisch, L.B. Schorr and C.H. Weiss, eds., *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts*. Washington: The Aspen Institute, pp.65–92.

Weiss, C.H., 1997a. How Can Theory-Based Evaluation Make Greater Headway? *Evaluation Review*, 21(4), pp.501–524.

Weiss, C.H., 1997b. Theory-Based Evaluation: Past, Present, and Future. *New Directions for Evaluation*, 76, pp.41–55.

Weiss, C.H., 1998. Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), pp.21–33.

Wholey, J.., 1987. Evaluability Assessment: Developing Program Theory. *New Directions for Evaluation*, Spring 198(33), pp.77–92.

Wimbush, E., Montague, S. and Mulherin, T., 2012. Applications of contribution analysis to outcome planning and impact evaluation. *Evaluation*, 18(3), pp.310–329.

Winnicott, D., 1989. Transitional Objects and Transitional Phenomenon. In:

Playing and Reality. London: Routledge.

Wood, S. and Mayo-Wilson, E., 2011. School-Based Mentoring for Adolescents: A Systematic Review and Meta-Analysis. *Research on Social Work Practice*, pp.1–13.

APPENDICES

Appendix 1 Summary of approaches to using programme theory

	Approach	Main author(s)	Key aim/purpose	Key component(s) of programme theory	Advantages	Drawbacks
Testing programme theory	*Theory-driven evaluation*	Chen; Rossi	Anticipate programme effects and use social science theory to collect and analyse data to assess the extent to which theories hold	Social science knowledge; a priori knowledge	Avoidance of goal-trap; increased validity; ability to generalise	Lack of social science theory in many cases; programmes are not always defined and bounded entities a priori; difficult to quantitatively test theory; cost and time intensive nature
	Theory-based evaluation	Weiss; Aspen Institute	Evaluate the impact of comprehensive community initiatives	Theory-of-change (pathway of change, indicators of preconditions to outcomes, interventions used to bring about preconditions, other assumptions)	Programme lifecycle approach (from planning to evaluation)	Difficult to elicit all elements of theory-of-change due to lack of evidence; 'diluted' approach in practical guidelines
	Realistic evaluation	Pawson; Tilley	Identifying key context-mechanism-outcome (CMO) configurations across programmes to generate transferable lessons through testing and refining middle range theories	Middle range theory (programme theory - between day-to-day operations and social science theory)	Realist philosophy; generative causation	Difficulty in implementing realist philosophy of science
	PTDES	Donaldson	Formulate & prioritise evaluation questions and answer those questions (similar to theory-driven evaluation)	Social science knowledge; stakeholder knowledge	Simple 3-step approach; participatory	Finding a balance between complexity and usability

	Logic analysis	Brousselle; Champagne	To test the plausibility of a programme's theory using scientific knowledge prior to investing time/money in evaluation through identification of critical conditions/consideration of alternative programmes	Conceptual (social science knowledge/social science theory) & empirical (empirical observations)	Prior to investing time/money in evaluation; simple 3-step process; relatively quick as compared with evaluation	Little guidance on the extent to which scientific knowledge should be consulted
Other purposes	*Process/ implementation evaluation*	Patton (but in earlier ideas)	To consider what occurred in the programme (understanding of what actually happened in a programme) and analysis of how things happen i.e. the internal dynamics of the programme (process)	Empirical observation of the programme	Can readily identify where improvements can be made	Can be carried out in various ways; does not necessarily address outcomes
	Contribution analysis	Mayne	To utilise performance measures to highlight the contribution of a programme to the intended outcomes rather than the assumed need for a full evaluation	Performance measures to create a 'credible' picture of attribution through analysis of the problem & other factors at play outside the programme	Using performance measures, it is more continual rather than a stand-alone evaluation; less costly, time intensive and technical than other approaches	Contribution not attribution; levels of uncertainty
	Evaluability assessment	Wholey	To ensure a programme is properly defined and implemented prior to summative evaluation	Programme theory as part of the process	Done prior to investing in evaluation; inform programme improvement; participatory	Programme theory not always included & so often neglected

Appendix 2 Qualitative interview study – question guide

THIRD SECTOR ORGANISATIONS

QUESTION GUIDE

This study is concerned with evaluation practice in small third sector organisations,, and in particular the use of qualitative frameworks of programme theory (anything which represents, illustrates, describes the assumptions underlying the outcomes a programme aims to achieve, the ways in which those outcomes might come about, as well as how the programme aims to change those outcomes). However, this interview will focus primarily on your experiences of conducting programme evaluation more generally.

More information about this study is available in the 'Participant Information Sheet' which has been made available to you prior to conducting the interview.

[Confirm/check participant understood & has signed consent form]

QUESTION GUIDE

1. Tell me a little bit about the programme & your organisation

2. Their role (w.r.t the programmes & evaluation).
    a. What does the evaluation team look like?
    b. Who is involved?

3. Evaluative activity – 'walk me through your evaluation process'
    a. What does this involve?
    b. Who does evaluation?
    c. Conceptualising outcomes and how to get there: how do you know the project is making a difference? What is that knowledge based on?
    d. Knowledge and use of elements of programme theory?
    e. Diagrams/documents/frameworks – how do you develop these? Can you show me?

4. What are the skills they have to do that?
    a. What skills do they lack?
    b. Enablers and barriers of evaluation?

5. Evaluation use:
    a. How do you use evaluation?
    b. For what purposes?
    c. What do you use?
    d. Have you had any feedback from evaluations?
    e. What have you learnt from evaluation findings in the past?
    f. What have you learnt from the evaluation process in the past?
    g. How this learning is/can be better facilitated?

6. What are your thoughts on evaluation as a whole? How could it be better?

7. Do you work with any other projects/organisations w.r.t evaluation?

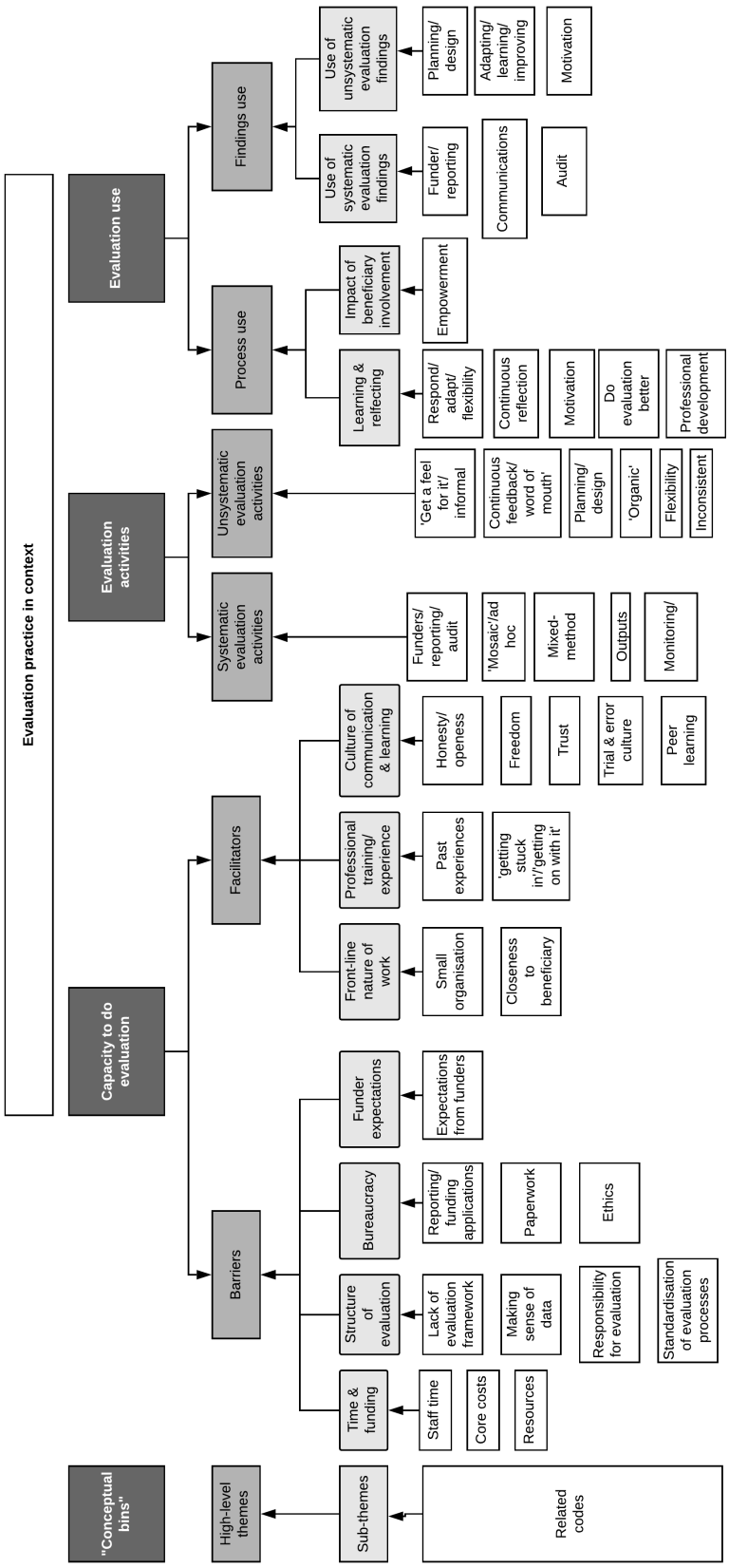
8. Are there any influences on how you do evaluation?

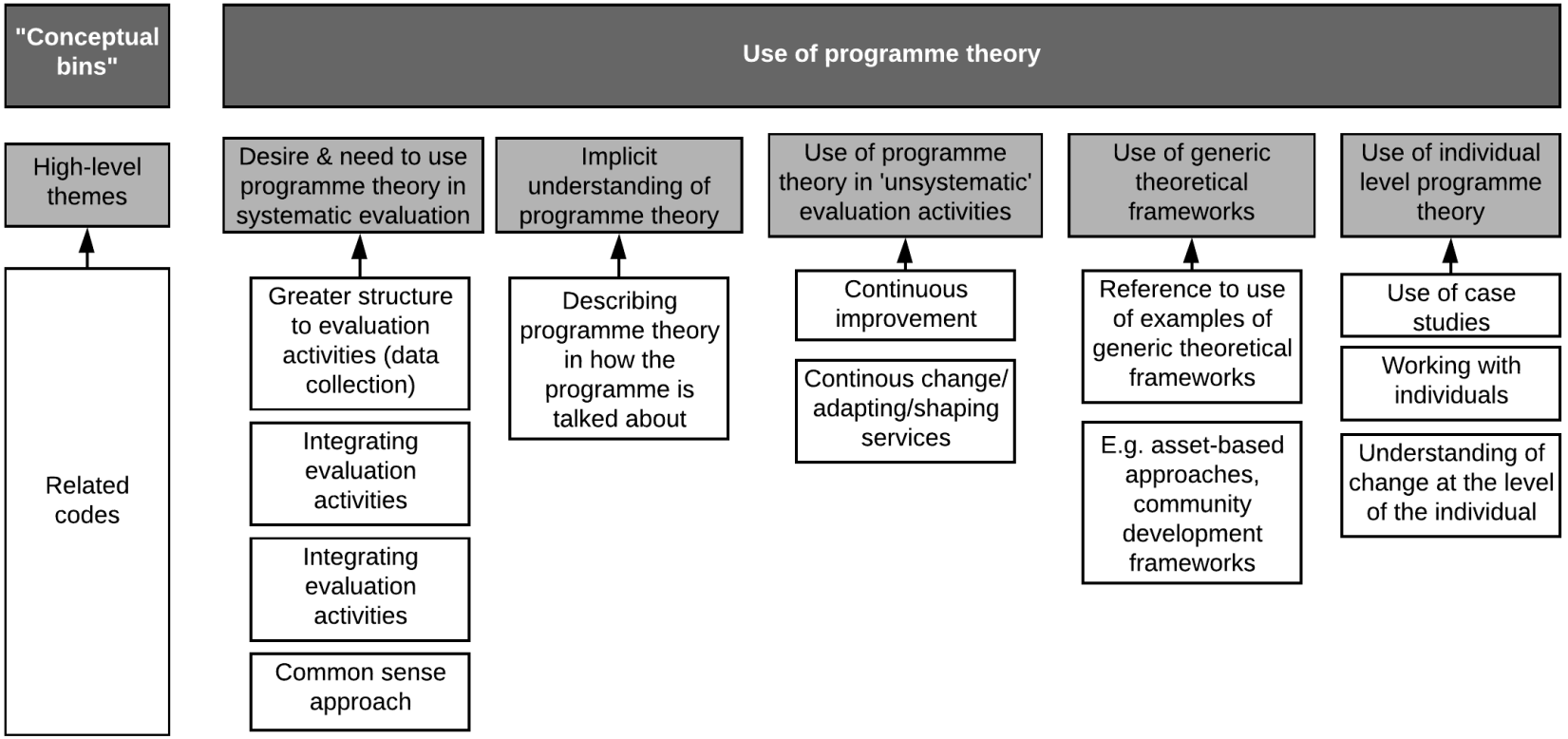
9. Can you tell me a bit about how you plan for and design the work that you do here?

As that is all of the questions I would like to address, I would like to give you the opportunity to ask me any questions or address any further points you think are relevant to this interview?

[If nothing] *Thank-you once more for your time. Please do not hesitate to get in touch with me should you have any other questions.*

Appendix 3 Study 1 coding hierarchy




"Conceptual bins"
Use of programme theory
High-level themes
Related codes
Desire & need to use programme theory in systematic evaluation
Implicit understanding of programme theory
Use of programme theory in 'unsystematic' evaluation activities
Use of generic theoretical frameworks
Use of individual level programme theory
Greater structure to evaluation activities (data collection)
Integrating evaluation activities
Integrating evaluation activities
Common sense approach
Describing programme theory in how the programme is talked about
Continuous improvement
Continous change/ adapting/shaping services
Reference to use of examples of generic theoretical frameworks
E.g. asset-based approaches, community development frameworks
Use of case studies
Working with individuals
Understanding of change at the level of the individual

Appendix 4 Organisation 1 question guide

ORGANISATION 1

QUESTION GUIDE

Thank-you very much for speaking to me today. The reason I'm doing these interviews is so we can find out a bit more about how you think this programme works for you. Using the information from these interviews I will create 'maps' which will make a visual diagram of what you all tell me. We will then use these to help make the programme better for other young people. Does that sound ok to you?

[Check participant has understood participant information sheet & consent form]

Tell me a little bit about yourself [opening questions]

i. How are you today?
ii. How is the ambassador training going? What do you do there?
iii. How is school? Do you enjoy it?
iv. What is your favourite subject? What do you least like about school?

Tell me about mentoring.

i. How do you describe to your friends/family what mentoring is or what you do with your mentor?
ii. What do you do in mentoring sessions?
    a. (PROBE) *What does that involve? What does this help you with? How does this help you?*
iii. How does mentoring make you feel? What is it that makes you feel that way?
iv. What about outside of mentoring? Do you use what you have done/spoken about in mentoring in school/outside of school?
v. You're an ambassador, tell me about that?
    a. (PROBE) Why did you decide to do that?
    b. What does that involve?
    c. How does this make you feel?

Why do you [*think you*] do mentoring?

i. When they gave you the option of having a mentor & being part of the programme, what made you decide to do it?
ii. You said you wanted to do XXX when you leave school. Does mentoring help you with that? How?
iii. For you, what is the best thing about mentoring?
    a. (PROBE) Why? Is there anything you don't like or that you wish was different?
    b. What do you want to achieve with mentoring? What do you want to get out of it?
iv. Are there any other things that help you achieve this?

Why it works or not for them? Other contextual factors

i. Is there anything which helps make mentoring work better?
ii. Is there anything which makes mentoring difficult?
iii. Tell me about your mentor? (if not already discussed)
iv. How is your mentor different from:
    (a) A friend?
    (b) A teacher?

School/friends/family and mentoring Social factors

i. What do your friends or family think about mentoring?
    a. (PROBE) Do you think this helps? How do you think this helps?
ii. Do any of your friends do mentoring too?
iii. Why doesn't everyone do mentoring?
iv. Is there anything you would change about mentoring or the programme in general?
v. What do you think life would have been like without mentoring or the programme in general?

A big thank-you! I have no more questions; do you have anything else you want to tell me or ask me? If not, I'd like to thank-you once again for your time!

Appendix 5 Organisation 1 participant information sheet

Project Information

University of Strathclyde Business School

[redacted]

My name is Robyn Millar and I am a research student
Department of Management Science
University of Strathclyde
robyn.millar@strath.ac.uk

What is this project looking at?

I am interested in the mentoring project you are part of.

I am interested in how you think the project works for **you** & how it helps you.

This information will help to understand how [redacted]an work better for everyone in the future.

What will happen?

I would like to chat with you about how you think [redacted]entoring works for you & how it helps you.

Together we wi[redacted].

The best part is there are no right or wrong answers ☺

Do I have to take part?

No! This is your choice to take part. Even if you do decide to take part, you can decide you do not want to anymore at any point!

What happens to my information?

If we speak together, I will take some notes, [redacted]ogether on paper, and I will record our conversation on a CD (not video). This is just in case I forget anything important!

I will keep all this information locked in my work cabinet & stored on computer password-protected storage.

I will never use your name – codes will be used.

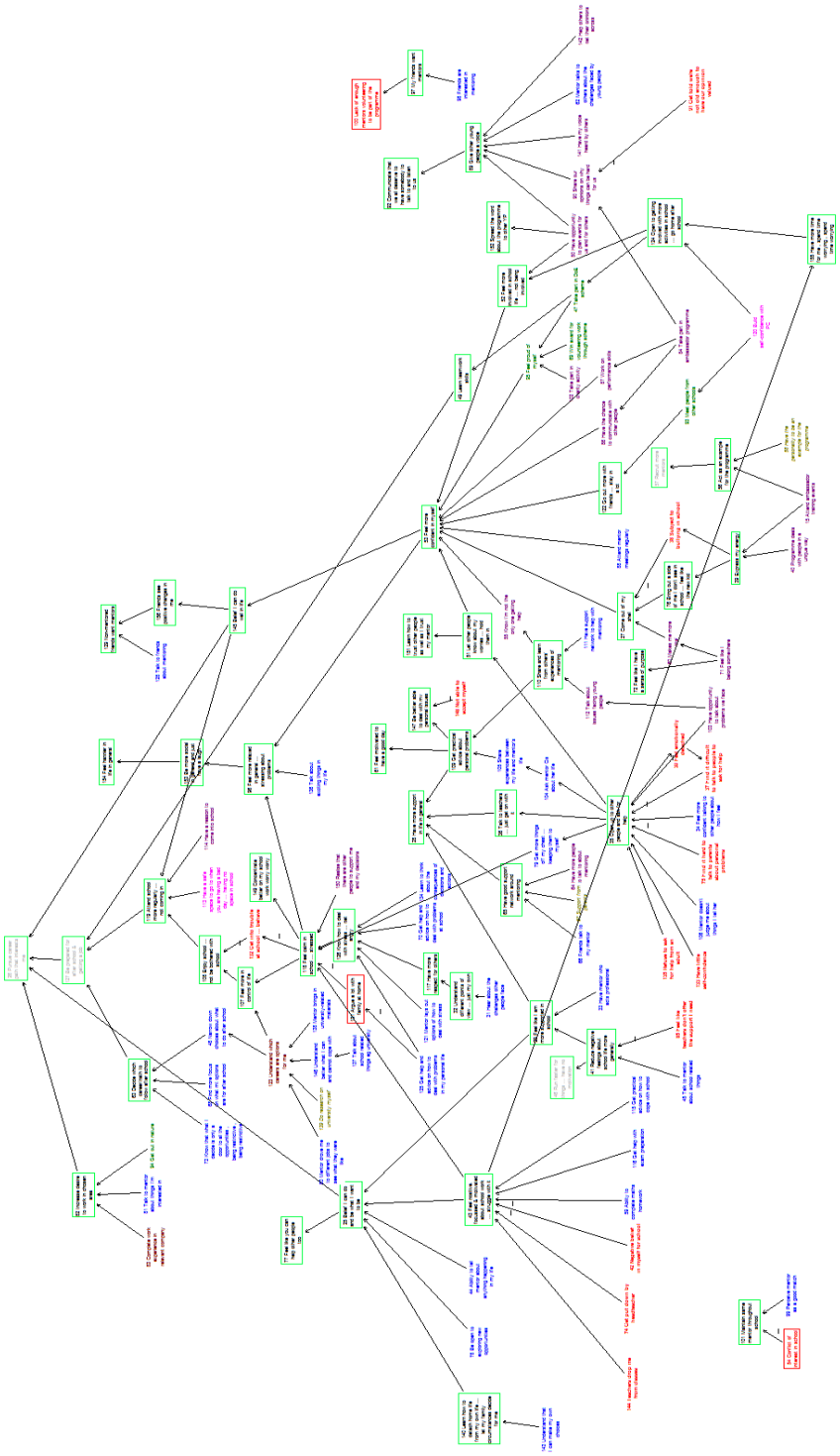
What happens next?

Talk to your school pathways co-ordinator & your parent/carer about this project. Let them know if there is anything you are unsure about.

If you agree, I will come and meet you & you can ask me any questions you have. I will then ask you to sign a form to say you would like to take part – you can still say no at any point! [redacted]ll also ask your parent/carer to sign this form. When I get that back, we will arrange a date and time to meet in your school to have our chat and map building session!

When the project is finished, I will write a report which will help YGT in the future. I will write a report for you to see too!

Appendix 6 Organisation 1 meta-outcome map



Appendix 7 Organisation 1 ‘decision to participate’ map