

Exploring evaluation using social practice theory in development
education: a longitudinal in-depth case study

Joana Zozimo

BA, MA (Hons)

July, 2016

Department of Educational Research,

Lancaster University, UK

Exploring evaluation using social practice theory in development
education: a longitudinal in-depth case study

Joana Zozimo

This thesis results entirely from my own work and has not been offered previously for
any other degree or diploma.

Signature



Abstract

This thesis explores how evaluation has been practised in a non-profit domain that of an English development education (DE) organisation. The study adds to explanations of how social practice theory contributes to the refinement of the understanding of evaluation practice. Generally, studies on evaluation practice remain unclear on how small, a-political non-profit organisations practice evaluation (Henry and Mark, 2003); especially, how their evaluation practice changes over time and with what effects (Saunders *et al.*, 2005). Within the specific non-profit sector of development education studies have described the insufficient knowledge of how these organisations practise evaluation (Bourn, 2014).

This is a single case study using an in-depth qualitative case study approach (Simons, 2009) to longitudinally explore evaluation practice and its change over time. Data from coordinators, practitioners and funders was ethnographically collected, during 1 year, through analysis of documentary evidence, 16 semi-structured interviews and 134 hours of observation where participants articulated their experiences of evaluation whilst identifying the influences shaping their evaluation practice. These articulations were coded and thematically analysed with findings displayed in a timeline of evaluation practice (Shove *et al.*, 2012). The empirical data obtained suggests that the participation of practitioners in evaluation appears to be taken-for-granted, but their full participation may be restricted. It also sheds light on the flow of power relationships that have traditionally been operated top-down, yet an internal layer of power within the non-profit leadership seems influential. The data analysed indicates that the co-occurrence of evaluation practice with other working activities shapes and alters how evaluation is practised in a development education setting.

Two ways this thesis extends the theory of evaluation practice are: first, it proposes that a social practice view of evaluation can enhance practitioners' experiences of evaluation by tracking how their practice has changed over time. This study makes an original theoretical contribution to the broader literature of evaluation practice informed by social practice theory (Saunders 2000; Saunders *et al.*, 2011), through the use of an advanced framework of the dynamics of social practice (Shove, 2009; Shove *et al.*, 2012), to explore the everyday life of evaluation and how it changes. Second, the thesis extends the theory of evaluation practice (Saunders *et al.*, 2011) by applying the dynamics of social practice into a novel research domain of development education evaluation. This thesis also offers an empirical contribution to the under-researched domain of development education (Bourn, 2014), by extending current knowledge on how evaluation is practised. Finally, the thesis contributes to advance the method of analysis of change over time, by having used a timeline as a tool to display findings, rather than only to collect and organise the dataset, as in other methods (Giele and Elder, 1998).

List of Contents

1	Introduction.....	17
1.1	Introducing the research	17
1.2	Justification for the research.....	18
1.2.1	The journey	19
1.2.2	Research context and the sample	21
1.2.3	Research problem, goals and questions	22
1.2.4	A longitudinal in-depth case study	23
1.2.5	Contribution	23
1.3	Overview of the thesis	24
1.3.1	Theoretical positioning	24
1.3.2	Methodological approach.....	25
1.3.3	Structure of the thesis.....	25
2	Theoretical foundations of evaluation practice.....	29
2.1	Introduction and structure of the chapter	29
2.2	Approach to the literature	30
2.3	Defining the interface between evaluation and practice.....	32
2.3.1	Evaluation	32
2.3.2	Practice.....	34
2.3.3	Evaluation Practice	35
2.4	Mapping the landscape of evaluation practice: theoretical perspectives.....	36
2.5	Current research debates in evaluation practice	48
2.5.1	The influence of context in evaluation practice	48
2.5.2	Non-profits organisations and performance.....	53
2.5.3	Power relationships in evaluation practice	59
2.5.4	The role of practitioner in evaluation practice	75
2.6	Summary and conclusion	77
3	Theories of practice: a social practice perspective.....	80
3.1	Introduction and structure of the chapter	80
3.2	Origins and contextualisation	82
3.3	The elements of practice	84
3.4	Trajectories of practice.....	88
3.4.1	Practitioners as carriers of practices	89
3.4.2	Practitioners' encounters: communities of practice.....	91
3.5	The co-occurrence of practices.....	94
3.5.1	The life paths and daily paths of practitioners	95
3.5.2	The formation of practices	97
3.5.3	The effects of the co-occurrence of practices	102
3.6	Summary and conclusion	104

4	Exploring a social practice perspective of evaluation.....	106
4.1	Introduction and structure of the chapter	106
4.2	Research using social practice theory.....	108
4.3	Applications to evaluation practice	109
4.4	Formulation of the research questions.....	116
4.4.1	How do Development Education organisations evaluate their social interventions?.....	119
4.4.2	What are the Northwest Development Education Centre stakeholders doing when they are evaluating the youth project?.....	120
4.4.3	What are the influences that shape the Northwest Development Education Centre’s evaluation practice, particularly the youth project?	121
4.4.4	How has evaluation practice changed across the youth project’s evaluation timeline?.....	121
4.5	Conclusions	123
5	Methodology	126
5.1	Introduction and structure of the chapter	126
5.2	Research design	128
5.2.1	Research philosophies, epistemological assumptions and methodological approach.....	128
5.2.2	Social constructionist approach adopted.....	130
5.2.3	Justifying the qualitative design.....	131
5.2.4	Case study research design	133
5.3	Approach to fieldwork.....	136
5.3.1	Situating the research and the researcher in context.....	136
5.3.2	Social practice theoretical framework employed.....	138
5.3.3	Research setting	145
5.3.4	Sampling procedures.....	147
5.3.5	A composite profile of the research participants	148
5.3.6	Data collection methods.....	155
5.3.7	Data analysis and interpretation.....	166
5.3.8	Ensuring data quality	182
5.3.9	Ethical considerations in the fieldwork.....	187
5.4	Reflecting on the research journey	188
5.4.1	Methodological, technical and other issues encountered.....	188
5.5	Summary and conclusions to the chapter	195
6	Participation in evaluation practice	198
6.1	Introduction and structure of the following chapters.....	198
6.2	The context of the youth project: background information.....	206
6.2.1	Funding dependency versus competence.....	208
6.3	Barriers to practitioners’ participation.....	212
6.3.1	Leadership.....	212
6.3.2	Funding	214
6.3.3	Co-occurrence of working activities.....	215

6.4	The typology of participants' engagement in the practice of evaluation	218
6.4.1	Disengagement.....	218
6.4.2	Forcing	221
6.4.3	Resistance	223
6.4.4	Coping.....	226
6.4.5	Collaboration.....	229
6.5	The change observed in evaluation practice: the paradox of participation.....	233
6.6	Summary and conclusions to the chapter	234
7	Power relationships in the evaluation practice	237
7.1	Introduction to the chapter	237
7.2	The context of a funder-recipient based evaluation: balanced and imbalanced-power relationship	239
7.3	Configurations of power relationships across the evaluation timeline.....	246
7.3.1	Leadership-led power relationship.....	246
7.3.2	Expected-led power relationship.....	253
7.3.3	Competence-led power relationship	256
7.4	The influence of power relationships in the practice of evaluation: the role of perception	259
7.4.1	Funders and Coordinators	259
7.4.2	Coordinators and Practitioners.....	261
7.4.3	Practitioners amongst themselves	263
7.5	The consequences of power relationships for evaluation in a Development Education context.....	264
7.5.1	Legitimised dishonesty	265
7.5.2	Judgement in evaluation	268
7.5.3	Scepticism in evaluation: a sceptical-resilient practitioner.....	273
7.6	The change observed in evaluation practice: modification /stabilisation and competition	279
7.7	Summary and conclusions to the chapter	280
8	The co-occurrence of evaluation practice with other working activities	285
8.1	Introduction to the chapter	285
8.2	Evaluation practice and working activities in a Development Education workplace.....	287
8.2.1	Evaluation practice.....	287
8.2.2	Working activities	289
8.3	The co-occurrence of practices through practice-change episodes	293
8.3.1	Different set of expectations in funder's evaluation approach	294
8.3.2	An extraordinary meeting	299
8.3.3	NDEC's change of office.....	304
8.4	The change observed in evaluation practice	310
8.4.1	Overlapping.....	310
8.4.2	Dominance	313
8.5	Summary and conclusions to the chapter	314

9	Conclusions.....	317
9.1	Introduction and structure of the chapter	317
9.2	Revisiting research goals and answering research questions.....	319
9.3	Contributions of the study.....	328
9.3.1	Contribution to theory of evaluation practice	328
9.3.2	Contribution to practice of development education.....	332
9.3.3	Contribution to method.....	335
9.4	Implications for future research	335
9.4.1	Implications for the domain of evaluation practice	335
9.4.2	Implications for the domain of development education	337
9.4.3	Implications for practitioners in the development education domain.....	338
9.4.4	Implications for other practitioners.....	338
9.5	Limitations and suggestions for future research.....	339
9.5.1	Limitations of the research.....	339
9.5.2	Suggestions for further research	340
9.6	The journey revisited: the “writing retreat” as a daily practice.....	342
10	List of references	345

Acknowledgements

This thesis is dedicated to my amazing husband, Ricardo Zozimo, for his enthusiasm, endless support and happiness, throughout my PhD journey. He is the one who has always believed that I could and will do it, encouraging me to apply for this PhD on the day before Manuel was born, five years ago. I definitely couldn't have done it without you! We are such a good team!

I also offer this work to my sons Francisco and Manuel, who together with Ricardo, made this endeavour possible through cooking, cleaning, and cheering me up along the way. This project is yours and hope one day you will read it!

I would like to extend a heartfelt thanks to my mother and grandmother, who have unconditionally supported me and given the best they knew for my education.

I am forever thankful to my former supervisor and mentor, Ulrich, who tried to convince long time ago that this was the route to follow. You were responsible for this!

I am sincerely grateful to my supervisor, Dr. Ann-Marie Houghton. I know it would have been not possible without your patience, support and invaluable guidance with this 'Portuguese stubborn student'.

I offer my thanks to Lancaster University Peel Studentship Award and to William Ritchie and Friend's Programme, which have made possible the conclusion of this thesis.

This thesis could not have happened without the virtual inspiration of Brian, Esther, and Brene; the support of Chryssie, who taught me how to breathe in stressful periods and the friendship of my extended family in Lancaster, Virginie, Sheila, Sandra, Sue, Armineh, Steve, Ceu, Joe, Rosalba, Sole, Marcela, Elena, Ana, Patrick, Telma, Mafalda, Carol, CLC, Dee, Matt, and Jo. Thank you all for being there, for keeping my energy levels up and for having looked after the boys when most needed.

A warm thanks to unique persons in my life such as, Grupeta, Sara, Augusto, and Ana Maria.

Finally, I offer my thanks to the individuals that have participated in this research. This thesis could not have happened without their enthusiasm and openness towards this investigation.

List of abbreviations

CLPR	Competence-Led Power Relationships
COoEP WAs	Co-occurrence of Evaluation Practice with other Working Activities
COoEP WAs-YPET	Co-occurrence of Evaluation Practice with other Working Activities in the Youth Project Evaluation Timeline
DAF	Development Awareness Fund
DE	Development Education
DEC	Education Centre
DE NPO	Education Non-Profit Organisation
DfID	Department for International Development
DSP	Dynamics of Social Practice
EP	Evaluation Practice
EC	European Commission
ELPR	Expected-Led Power Relationships
GMA	Grant Management Agency
LLPR	Leadership-Led Power Relationships
M&E	Monitoring and Evaluation
NDEC	Northwest Development Education Centre
NPOs	Non-Profit Organisation
PAR-YPET	Participation in the youth project evaluation timeline
PCEs	Practice-Change Episodes
PET	Program Evaluation Theory

PR	Power Relationships
PRE	Power Relationship Episodes
PR-YPET	Power relationships in the youth project evaluation timeline
SP	Social Practice
SPT	Social Practice Theory
VPA	Values-Pluralism Approach
YP	Youth Project
YPCT	Youth Project Chronological Timeline
YPEP	Youth Project Evaluation Practice
YPET	Youth Project Evaluation Timeline
YPP	Youth Project Partnership
YW	Youth Work

List of Figures

Figure 2.1 Structure of Chapter Two	30
Figure 3.1: Structure of Chapter Three	81
Figure 4.1: Structure of Chapter Four	107
Figure 5.1: Structure of Chapter Five	127
Figure 5.2: The order of research elements (adapted from Crotty, 1998)	129
Figure 5.3: Multiple points and sources of data collection per stage.....	142
Figure 5.4: Multiple sources of data per alphabetic order	143
Figure 5.5: Phases of data collection	159
Figure 5.6: Network view of thematic analysis in ATLAS ti.....	173
Figure 5.7: Network view refinement in ATLAS ti.....	176
Figure 5.8:List of memos written during the analysis in ATLAS ti.....	178
Figure 6.1: Youth project evaluation timeline: chronological facts	203
Figure 6.2: Participation in the youth project evaluation timeline (PAR –YPET)	205
Figure 7.1: Power relationships in the youth project evaluation timeline (PR-YPET)	238
Figure 7.2: Balanced and imbalanced power-across the YPET.....	243
Figure 7.3: Configurations of power relationships across YPET	246
Figure 7.4: Power relationships shaping evaluation practice through perception	259
Figure 7.5: Consequences of power relations shaping evaluation practice in a DE setting ..	264
Figure 8.1: Co-occurrence of evaluation practice with other working activities in the youth project evaluation timeline (COoEP with other WAs-YPET).....	286
Figure 9.1: Structure of Chapter Nine.....	318
Figure 9.2: Research questions answered	321
Figure 9.3 Overview of how evaluation practice changed across the YPET.....	327

List of Tables

Table 2.1: Chronological list of the main studies linking evaluation theory and practice.....	39
Table 2.2: Main theoretical perspectives to research evaluation practice	42
Table 2.3: Stages of the cycle of the project (EuropeAID, 2004, p.33).....	52
Table 2.4: Features of non-profit leadership (source: Alaimo, 2008).....	65
Table 3.1: Frameworks of elements of practice by Reckwitz (2002) and Shove <i>et al.</i> (2012)	85
Table 4.2: Overview of the research questions and their rationale.....	118
Table 5.1: Embedded case study (source: Yin, 2008)	135
Table 5.2: Composite profile of a DE coordinator	149
Table 5.3: Composite profile of a DE practitioner	151
Table 5.4: Composite profile of a DE funder	153
Table 5.5: Preferred methods in case study research (adopted from Simons, 2009).....	156
Table 5.6: Semi-structured interviews	159
Table 5.7: Responsive interview type of questions (Rubin and Rubin, 2012)	161
Table 5.8: Observation sessions.....	163
Table 5.9: Documentary evidence	164
Table 5.10: Categorisation of documentary evidence.....	165
Table 5.11: Overview of the data collection phases per data sources	165
Table 5.12: Informal conversations and extracts collected.....	166
Table 5.13: Influential approaches to my analysis and interpretation (source: author).....	168
Table 5.14: Preparing analytical data in ATLAS ti.	171
Table 5.15: Example of analytical coding in ATLAS ti.	174
Table 5.16: Methodological and technical issues encountered.....	189
Table 5.17: Summary of methodological decisions.....	197

Table 6.1: Practice-change episodes discussed.....	199
Table 6.2: Research findings and issues emerged	202
Table 6.3: Summary of findings on participation (PAR).....	235
Table 7.1: Features of non-profit leadership in a DE setting (adapted from Alaimo, 2008).250	
Table 7.2: Summary of findings on power relationships (PR)	281
Table 8.1: Typology of evaluation practice (adapted from Carman, 2007).....	288
Table 8.2: Typology of working activities in a development education workplace (adapted from Carman, 2007).....	290
Table 8.3: Practice-change episodes in a Development Education working environment....	294
Table 8.4: Categorisation of overlapping in a co-occurrence working environment	311
Table 8.5: Summary of findings on the co-occurrence of evaluation practice with other working activities.....	315

List of Extracts

Extract 5.1: Extract from my learning journal, December 2012.....	192
Extract 5.2: Extract from my learning journal, October 2012	194
Extract 5.3: Extract from my learning journal, January 2013.....	195
Extract 6.1: Extract from field notes, August 2012	229
Extract 7.1: Extract from field notes, May 2012	241
Extract 7.2: Extract from my learning journal, November 2012	257
Extract 8.1: Extract from field notes, November 2012.....	299
Extract 8.2: Extract from my learning journal, November 2012	301
Extract 8.3 Extract from my learning journal, January 2013.....	304
Extract 8.4 Extract from my learning journal, March 2013.....	305

List of Appendixes

Appendix 1: Mind map with sample of the literature review.....	360
Appendix 2: Development Education.....	361
Appendix 3: Glossary.....	363
Appendix 4: Data collection.....	364
Appendix 5: Evaluation form	370
Appendix 6: Reflective boxes	371
Appendix 7: Data analysis and interpretation.....	373

1 INTRODUCTION

This chapter introduces the study by providing a summary of the journey and the research context, problem and questions that stems from a critical review of the literature. In addition, it illustrates how the contribution made by this thesis is linked to specific research gaps found in the literature. The chapter introduces the theoretical orientation and methodology employed in the study, and finalises with an outline of the structure of the thesis.

1.1 Introducing the research

This study deepens current understanding of how evaluation is practised in a specific non-profit setting of Development Education (DE). Current theoretical works have explored the evaluation practice (EP) within the non-profit sector, in general (Rossi *et al.*, 2005; Morra-Imas and Rist, 2009; Carman and Fredericks, 2010); however, previous studies have reported that the non-profit sector “*knows very little about how evaluation is being practiced, by whom and where it is being practiced and to what effect*” (Henry and Mark, 2003, p.69). Early conceptualisations about EP have largely theorised drawing upon performance-based frameworks which emphasised results-oriented evidence, certainty, and attribution (Kusek and Rist, 2004; Carman, 2007; King, 2007), rather than practice-based frameworks, which tend to focus on reflection, uncertainty, and contribution (Schön, 1983; Saunders *et al.*, 2005; Mayne, 2008; Patton, 2010). Whilst there is a reasonable amount of literature about the way non-profit organisations (NPOs) respond to the evaluation demand (Greene, 1999; Henry and Mark, 2003; Rogers, 2008; Thomson, 2010), the published studies are mostly quantitative or mixed-methods (Carman, 2007; Carman and Fredericks, 2010). In contrast, this study qualitatively explored how evaluation was practised in a DE setting through the investigation of a single in-depth case study –a youth project’s evaluation practice (YPEP). By

investigating this evaluation over time, the study also explored the trajectory of change of its evaluation practice (EP), through the thematic analysis of three groups of participants' accounts: coordinators, practitioners and funders. The focus of this study was an inductive effort to extend the overall knowledge of the EP within the DE sector; also informed by previous research on practice-based approaches to evaluation (McCluskey, 2011; Saunders *et al.*, 2011).

A distinctive feature of the research draws on a specific strand of social practice theory – the dynamics of social practice (Shove *et al.*, 2012) to examine how the YPEP changed over time; what participants did when evaluating the project - their “*routinised practice*” (Saunders *et al.*, 2011), alongside their other working activities, which may or may not concur; and what the influences were for the change of EP over time.

1.2 Justification for the research

As Shaw *et al.* (2006) noted “*research on evaluation will not be a magic bullet. [...] It will not replace all judgements about wise evaluation practice – but instead can aid such judgement to a significant extent*” (p.18). This research is important for four reasons:

1. It extends the literature that claims a performance-based approach is inadequate to evaluate social interventions, by researching how non-profit organisations (NPOs) evaluate their interventions (Greene, 1999).
2. It enables novel insights to emerge regarding the configurations of power relationships, by exploring why NPOs are perceived to inadequately evaluate their interventions (Carman, 2007).

3. It theoretically contributes to the EP domain, by exploring the potentialities of the dynamics of the social practice strand in researching evaluation in a DE domain (Shove *et al.*, 2012).
4. It critiques a performance-based approach to evaluating NPOs' social interventions, required by most funders, and proposes a social practice approach to evaluation (Saunders *et al.* 2011).
5. Personally, this research is important to answer challenges that emerged from my previous experience in the global world, which I briefly introduce next.

1.2.1 The journey

My interest in alternative ways to practise evaluation is not recent and the concept of evaluation as a social practice has been in my mind for a while. The idea that what one does in a routinised practice can be seen and valued as evaluation struck me the first time I read about it. This interest was deep-rooted in my background as a monitoring and evaluation (M&E) program officer, mainly with NPOs in African¹ and European² countries. In the last ten years I have seen various evaluation experiences based on a variety of perceptions and, as a result, my curiosity to expand the understanding of the EP has increased.

In 2010, I was privileged to collaborate with some NPOs in the UK, in a specific domain of DE, where I understood that for them evaluation had a specific performance-based meaning, driven by the funders. From this experience, I deeply reflected about two of my previous evaluation experiences; first, as an M&E practitioner in Mozambique, and second, as a consultant supporting NPOs in an M&E domain in Portugal.

¹ Cape Verde; Guinea-Bissau; Angola; Senegal; Gambia; Morocco; Algeria and Mozambique.

² Portugal, Belgium, Spain, and the UK.

In Mozambique, I worked for a Spanish NPO and participated, as one of the project managers, in an intermediate evaluation. At that time, I consciously observed people struggling with the term ‘evaluation’, and with the meaning assigned to it. Initially, I thought that their struggling was associated with different cultural approaches to evaluation, from a European funder perspective to an African partner, grassroots-based perspective. However, as part of the evaluation planning, I sat in on various funders’ meetings (in an advisory role) and began to track different reactions to evaluation, taking into account the diverse context in Mozambique. During these meetings, I also witnessed the social problems that some local communities live with; what was a priority for their survival (the need to feed their families) was not for the funders, and what was relevant for the funders (the practice of evaluation) appeared irrelevant for the local communities.

The way individuals, for example in Mozambique, struggled with evaluation made me relate to the possible reasons with their cultural context, but when returning to Portugal and later in the UK, I surprisingly, found the same challenge among practitioners. This was especially revelatory when I first volunteered in the Northwest Development Education Centre (NDEC), because they were about to plan a three-year evaluation that I later had the opportunity to longitudinally follow, as this thesis’ case study. Whilst I was reflecting on my previous world-wide experience of EP (with other types of complexities such as, gender-balance, poverty, and safety related issues), I noticed that in this ‘developed’ scenario the similar struggle with evaluation that I had seen in other countries, was present. From this discernment, I concluded that people struggled with evaluation, not necessarily, because of the cultural differences of the context, but possibly for intrinsic reasons of the phenomenon of evaluation itself. These experiences have shaped my ontological view of the meaning of reality and left the ‘seed’ of the research, as a resource to understand: the complexity of the social world, the place of evaluation to serve these inequalities, and my contribution as a

researcher (and global citizen). I was also interested in the negotiation taking place amongst these stakeholders' forums (how they promote one argument rather than another and how they build consensus and report evaluation for specific audiences). The lived observation of these practices was the bottom-line for my reflection and led to a profound interest in EP, mainly, in the particularities of each experience, each country and each community. Surprisingly, these observed experiences were already research skills, but I only fully discovered that later in the journey (Blaxter *et al.*, 2010).

The second experience was in Portugal, while I worked for a consultancy in supporting NPOs in the M&E domain. In this context, the challenge mostly shared by these organisations was the '*how do we know*' dilemma of having to produce performance indicators to identify if the projects have achieved the outcomes set. As a consultant, I used to apply the principles of program evaluation and theory of change, despite the need to translate them to the NPOs, but now that I have been introduced to a social practice view of evaluation, I am able to interpret the value of this experience and recall that time when I did not recognise that some NPOs' daily routines were already EP (see chapter three; theories of practice). These experiences, alongside the volunteering in the DE sector, have driven my personal interest and decision to research EP, in the form of a PhD.

1.2.2 Research context and the sample

Previous research in EP has emphasised the importance of selecting an appropriate research context in order to conduct these complex investigations (Patton, 2002). DE NPOs were selected as the main setting, informed by insight in the literature stating the relevance of research evaluation in the under-researched domain of DE (Bourn, 2014). The following sampling criteria was used: level of DE activities within intervention, receipt of public (or (private) funding; formal evaluation required by funding agency; located within the

Northwest of the UK. As a result, NDEC was purposively chosen as the case organisation and within that, the EP of a three-year project - the YP - was selected to be the single-case's unit of analysis (Yin, 2008; Simons, 2009).

1.2.3 Research problem, goals and questions

Past studies suggested that the practice of evaluation does not appear to have kept pace with the fast changing complex social world in which we live (Stevenson *et al.*, 2002; Newcomer *et al.*, 2004; Carman, 2007). The apparent standardisation of a performance-based approach to evaluation seemed problematic, both from the literature (Greene, 1999; Simons, 2003; Saunders *et al.*, 2011) and from my professional experience. For the purposes of accountability, the current research focused on the centrality of the role of the evaluation practices in responding to a complex, multifaceted and uncertain social world (Patton, 2010; Saunders *et al.*, 2011). Additionally, scant research on EP in the DE domain recognised challenges in evaluation of social interventions from a performance-based approach, calling for alternative responses to practise evaluation (Bourn, 2014).

Having identified some theoretical and analytical gaps³ in the literature on EP, this study aims to explore and gain further understanding on how DE organisations evaluate their social interventions. To accomplish this research goal one overarching question was formulated and narrowed into three sub-questions:

Overarching RQ: How do Development Education organisations evaluate their social interventions?

³ See Table 4.2; Overview of the research questions and their rationale.

RQ.1: What are the Northwest Development Education Centre’s stakeholders doing when they are evaluating the youth project?

RQ 2: What are the influences that shape the Northwest Development Education Centre’s evaluation practice, particularly the youth project?

RQ 3: How has evaluation practice changed across the youth project’s evaluation timeline?

1.2.4 A longitudinal in-depth case study

The YPEP was observed over twelve months, starting in March 2012, and data from coordinators, practitioners and funders was ethnographically collected in two phases. Twelve participants articulated their experiences of evaluation in retrospect, and identified the influences shaping their practice. Eighteen semi-structured interviews (including two pilot interviews with practitioners from different NPOs) were conducted in the first phase of data collection, during the implementation stage of the evaluation; and for purposes of analysing change over time; all twelve participants were invited for another interview, in the second phase of data collection, during the reporting of the evaluation. From those, four were interviewed a second time and probed regarding their shared experience of EP (see Table 5.3: Semi-structured interviews conducted). This longitudinal research design generated a rich data set that allowed for an in-depth exploration of EP in a DE context over time.

1.2.5 Contribution

In line with the latest theoretical developments of evaluation research, the original contribution of this thesis extends the theory of EP in two ways. Firstly, it extends the broader literature of EP informed by social practice theory (Saunders 2000; Saunders *et al.*, 2011), through the use of an advanced framework of the dynamics of social practice (Shove, 2009; Shove *et al.*, 2012), to explore the everyday life of evaluation and how it changes. Secondly,

it extends the theory of EP by applying the dynamics of a social practice strand (Shove *et al.*, 2012) into a novel research domain of DE evaluation. This thesis also offers an empirical contribution to the under-researched domain of DE (Bourn, 2014), by extending current knowledge on how evaluation is practised. Finally, the thesis contributes to advance the method of analysis of change over time, by having used a timeline as a tool to display findings, rather than only to collect and organise the dataset, as in other methods (e.g. life course analysis; (Giele and Elder, 1998)).

1.3 Overview of the thesis

This section provides an overview of the thesis, by presenting the theoretical and methodological orientations, and thesis structure.

1.3.1 Theoretical positioning

Recent theoretical developments in evaluation literature suggest that understanding practices is necessary to the process of evaluation, particularly when concerned with social change (Saunders *et al.*, 2011), such as DE (see Appendix 2a; background information on DE). This study builds upon the theoretical view of evaluation as a social practice (Saunders, 2012; Saunders *et al.*, 2011; Trowler *et al.*, 2012; Trowler, 2013) that draws on the initial work of past social practice theorists, to conceptualise evaluation as a routinised practice conducted on a daily basis (Giddens, 1979; Reckwitz, 2002; Schatzki *et al.*, 2001; Law, 2004).

This social practice view of evaluation may challenge and change how NPOs perceive the evaluation, in general, and their practice, in particular. In addition, the theoretical underpinning to explore the element of change across the trajectory of EP in the DE domain was informed by Shove *et al.* (2012), who advanced a framework for the analysis of the dynamics of social practice in everyday life (and how it changes). The theoretical aim of this

exploration is to apply Shove *et al.*'s lens, within a DE context, to probe the research and draw insights from my empirical data.

1.3.2 Methodological approach

The study draws on a social practice theoretical (SPT) lens to inductively explore how evaluation has been practised, and changed, over a three-year project. In methodological terms, this research builds upon a qualitative exploratory approach to get an in-depth account of how three groups of participants within the YP practised evaluation. The research also explores how EP has changed over time, and for this reason adopts a single in-depth case study, not only as a method but also as a research approach (Simons, 2009). This approach allows for a detailed understanding of a real-life phenomenon, that involves relevant contextual conditions highly appropriate to the phenomenon of EP within DE. The longitudinal dimension of the research combined with data collected ethnographically, through the embedded case design (Yin, 2008), allowed for the “*particularities*” of the single case to emerge (Simons, 2009, p.32). To generate new insights a “*thick-description*” (Geertz, 1973, p.6) of how evaluation has been undertaken is provided.

1.3.3 Structure of the thesis

This section provides a brief overview of the remaining chapters of the thesis.

Chapter two contextualises the study, by situating the thesis and its focus in the literature of EP. The chapter starts with a definition of the terms ‘evaluation’ and ‘practice’, and a description of how these intersect. It then critically maps the theoretical landscape of EP. As pointed out by scholars, existing knowledge has focused on the need for alternative approaches to evaluate social interventions, due to the inadequacies of performance-based

approaches in such a quest (Greene, 1999; Saunders *et al.*, 2011). Recognising the centrality of evaluation and its several theoretical roots, chapter two briefly describes the emergence of a social practice perspective in EP. The chapter also examines the current research debates in EP, particularly acknowledging a key debate for this research: DE NPOs and their performance evaluation practice.

Chapter three introduces social practice theory (Giddens, 1984; Shatzki, 1986; Reckwitz, 2002) as the guiding theoretical perspective of this thesis. The initial section introduces the origins and contextualisation of the theories of social practice illustrating how the specific strand of dynamics of social practice (DSP) have emerged (Shove *et al.*, 2012). This chapter explains the core messages of this theoretical strand, by reviewing relevant concepts for the analysis of evaluation as a social practice bound to change over time. These concepts are: the elements of practice; the trajectories of practice; practitioners as carriers of practice; the formation of practices; the co-occurrence of practices and the change in practices.

Chapter four explores how a social practice perspective of evaluation has been covered by past literature on EP and, in doing so, identifies the theoretical gap of this research. Despite the recognition of the need to explore change in EP (Saunders *et al.*, 2011), past studies have not yet employed Shove *et al.*'s (2003; 2007; 2012) lens of DSP to investigate, for example, concurrent practices and its effects for the practice of evaluation over time. Additionally, chapter four emphasises the potential of a DSP lens to examine how evaluation has been practised in a DE context. To critically discuss the intersection of DSP and EP, the chapter concludes with an outline of the gaps in which this intersection can benefit further research on EP.

Chapter five presents a detailed account of the methodology in three main sections. The first section provides an overview and justification of the research design adopted. The second section outlines the main research by explaining the approach to fieldwork, including the data collection methods used, the data analysis and the ethical aspects considered during the fieldwork. The third section describes the reflection on the research journey, by outlining some methodological and technical aspects encountered. It includes a 'reflective space' with ethically relevant stories from the fieldwork.

Chapter six is the first of three chapters that combine findings and discussion. It begins by introducing the structure of the second part of the thesis and provides the first timeline of the study – the YP evaluation timeline (YPET) with chronological facts. It then describes participants' experiences of participation in the YPEP over time. The chapter illustrates the participants' notion of 'forced participation', by exploring the barriers to participation, in a DE context. Further, the chapter discusses how participation influences participants' EP through a typology of their engagement in evaluation: disengagement, forcing, resistance, coping and collaboration in the practice of evaluation. The chapter includes a final section on the change observed in EP in relation to the theme of participation and the findings are visualised in a timeline for participation (PAR-YPET).

Chapter seven presents and discusses participants' accounts of power relationships in their experience of evaluation over time. The chapter examines how a funder-recipient based context is critical to perceive power relations and their balance and imbalance dynamics in a DE context. It also describes configurations of power relationships across the evaluation timeline, alongside how these influenced participants' practice of evaluation, through the role of perception. The chapter discusses how the configurations of power relationships shaped and changed participants' EP; and the consequences of these changes. The chapter concludes

by outlining the change observed in EP regarding the theme of power relationships, visualised in a timeline (PR-YPET).

Chapter eight examines and discusses the theme of co-occurrence in evaluation, by highlighting how the intersection of EP with other working activities (WAs) alters the former (the way evaluation is practised in a DE context). It begins by describing some practice-change episodes that illustrate some level of co-occurrence and then discusses the effect of overlapping and dominance between EP and other WAs and how it changed the former, over time. Findings on the theme of co-occurrence of EP with other WAs are visualised in a timeline (COoEP with other WAs -YPET).

Finally, chapter nine summarises how the three chapters of discussion answered this study's research questions. It revisits the research theoretical and empirical contribution, describing its contribution to knowledge, its implications for the domain of EP, for the domain of DE, for DE practitioners, and for practitioners in general. A final illustration combines previous timelines per finding to visualise the overview of how EP changed across the YP. The chapter also draws some suggestions for further research and finalises by revisiting the journey from a writer's practice perspective.

2 THEORETICAL FOUNDATIONS OF EVALUATION PRACTICE

2.1 Introduction and structure of the chapter

This chapter situates the thesis within the literature of evaluation practice (EP). It aims to define the boundaries of the research by providing the current understanding of the phenomenon of EP in non-profit organisations (NPOs).

The chapter starts by describing the approach taken to review the literature and the definition of EP used. It then outlines the theoretical landscape of EP, by identifying the main gaps in the literature. Next, four research debates are identified: first, the influence of context in EP; second, the contextualisation of NPOs and performance approaches where a specific review of development education (DE) is included; third, the emergence of power relationships in EP, and fourth, the role of practitioners in EP. The chapter finishes with a summary and the main conclusions from the review of the literature on EP. The overall structure is represented in Figure 2.1 (overleaf).

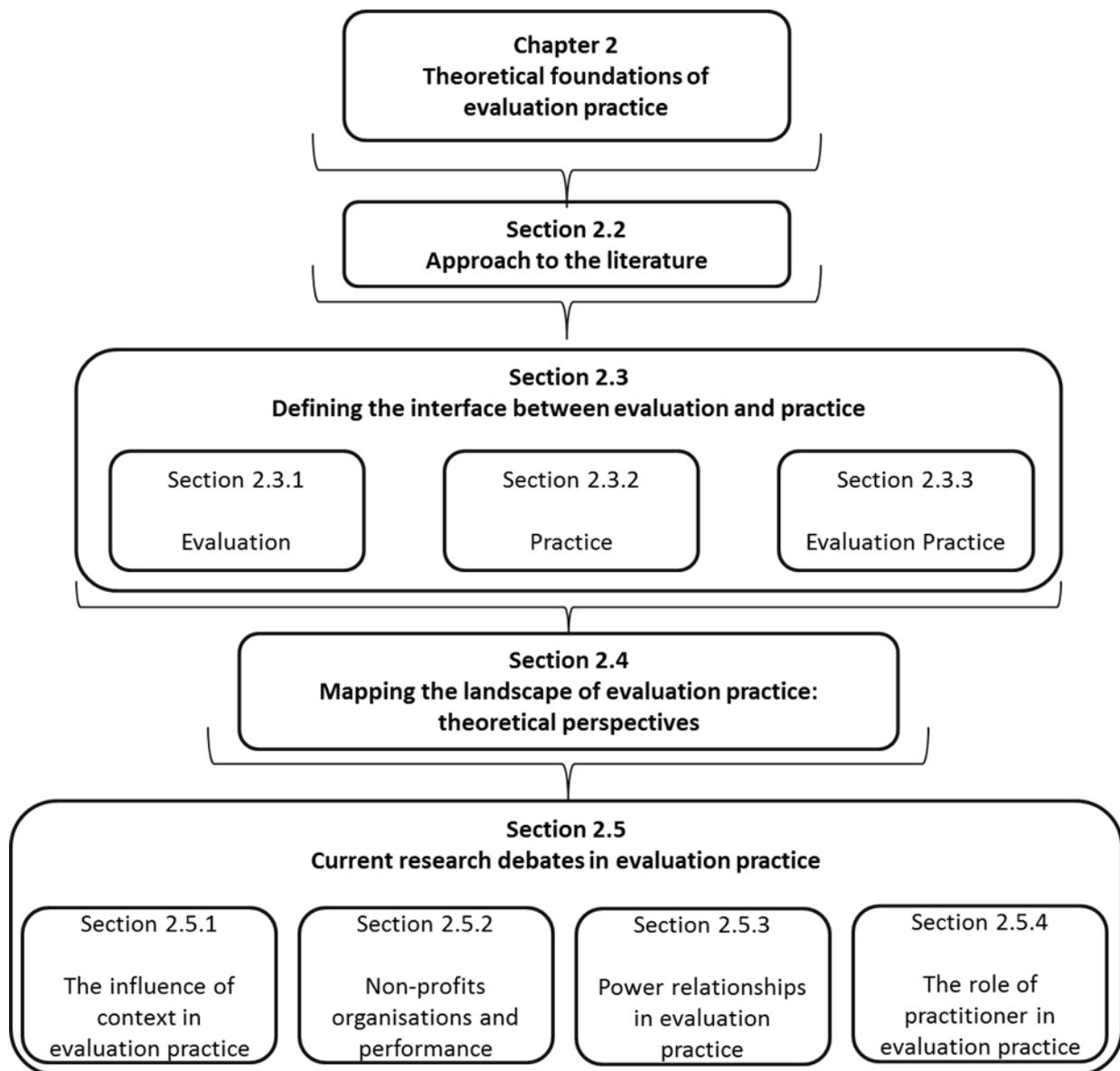


Figure 2.1 Structure of Chapter Two

2.2 Approach to the literature

I conducted a literature review in two periods throughout my research journey: the first period was at the beginning of the PhD, when I approached the literature to map the current knowledge about EP, particularly, the limitations and problems ‘known’. This was accomplished through the design of mind maps, through which I mapped the various strands of the literature on EP (see Appendix 1; mind map with sample of the literature review). I then narrowed my review to the relevant theoretical concepts that describe the present debate

of EP, alongside a detailed review of social practice theory (SPT), its origins, contextualisation, and principles. By closely examining the tradition of theorists whose views are applied to EP studies (Giddens, 1979; Schatzki *et al.*, 2001; Reckwitz, 2002; Schwandt, 2005; Saunders *et al.*, 2011), the review focused on how a SPT lens enhances the research of EP (Shove *et al.*, 2009, 2012; Saunders, 2012). A second period occurred during the writing-up of the thesis, when I contrasted the initial review of the literature with my findings; an important step forward in my research journey, because I clearly noticed the developments and contradictions that had arisen, as well as the key persisting themes.

The literature review presented, across chapter two to four, represents the EP domain, particularly, regarding how NPOs practise evaluation of social interventions. It includes previous studies from other domains of research, which have drawn on SPT applied to evaluation such as, e-learning (Bonamy *et al.*, 2004); social change and innovation (Westley *et al.*, 2006); Higher Education (HE) (Saunders *et al.*, 2011; Trowler *et al.*, 2012), and European social cohesion programmes (Saunders, 2011, 2012). Patterns, similarities and articulations across these domains pointed out theoretical gaps in the practice of evaluation; which informed my theoretical underpinnings, the generation of research questions and the research design.

The criterion for this review was threefold: firstly, to present previous research conducted in EP within NPOs in the social sector, specifically in a DE domain; secondly, the time period covered research from 1990 to 2015, because it was when key contributions on EP emerged, particularly, focused on a social constructivist view of evaluation (though - when appropriate, scholarly work outside this time frame was used); thirdly, to draw on previous research from theorists who have applied SPT to evaluation (Saunders *et al.*, 2011). I left relatively intact the review of DE as a discipline, because, for the purpose of this thesis, DE is the research

context. However, I examined some policy documents about evaluation (theory and practice) within a DE setting, to serve as background information. Taken together, the approach to the literature review outlined led to the identification of existing gaps and the formulation of research questions (See section 4.4).

2.3 Defining the interface between evaluation and practice

Evaluation practice (EP) is a widely recognised domain of evaluation that has recently seen a substantial increase in scholarly work (Alkin, 2003; Schwandt, 2005; Saunders *et al.*, 2011; Chelimsky, 2013; Chouinard, 2013; Leviton, 2015; Schwandt, 2015). To advance the discipline of EP, it is relevant to review its conceptual definition first, so I present the current understanding of both terms separately– evaluation and practice – to then examine them as a whole. Moreover, providing the individual definition of both terms is an appropriate task to explore the literature, as it allows the establishment of clearer research boundaries.

2.3.1 Evaluation

Evaluation has been defined in many ways and still is. According to the Oxford Dictionary, the origin of the word ‘evaluation’ links with the determination of the value of a project, program or policy (Simpson *et al.*, 2009). Evaluation has a variety of meanings that imply differing theoretical lenses, as well as purposes. For example, the OECD⁴ defines evaluation as: “*the **process** of determining the **worth** or significance of an activity, policy or program. It is as **systematic** and **objective as possible**, of a **planned, on-going or completed intervention**”* (OECD, 2002; **emphasis added**).

⁴ Organisation for Economic Co-operation and Development,
See http://www.oecd.org/document/32/0,3746,en_2649_33721_42632800_1_1_1_1,00.html#Evaluation; Accessed on 25th January 2013

Some authors argue that accountability purposes are the core issue in any evaluation (Huffman *et al.*, 2008; Carman, 2009), while others advocate learning as the main purpose (Riddell, 2001; Hoole and Patterson, 2008; McCluskey, 2011). For the former, a program evaluation theory corroborates accountability as key, whereas in the latter, a value-pluralism or social practice theory promotes the learning purpose of the evaluation. Some others argue that evaluations are needed either for improvement purposes (Vo and Christie, 2015) or to attribute results to a particular intervention, rather than to other potential causes (Kusek and Rist, 2004; Mayne, 2004; Morra-Imas and Rist, 2009). According to Scriven (1991), one of the founders of modern evaluation, there were nearly sixty different terms for evaluation, even 25 years ago. These include: “*adjudge, appraise, analyse, assess, critique, examine, grade, inspect, judge, rate, rank, review, score, study, test...*” (ibid. p.9). I would add that for each term, or group of terms, a specific theoretical approach applies, and for this reason it is relevant to map the previous research attempts to define evaluation according to a variety of theoretical approaches (see Table 2.1: chronological list of the main studies linking evaluation theory and practice).

A social practice (SP) dimension of evaluation has been widely accepted in the literature, in relation to evaluation of social interventions, because the purpose of evaluation in that dimension is to attribute the worth of how a specific intervention contributes to its overall goal; as such, the realisation that the process of attributing worth and value is social and relational has advanced the conceptualisation of EP. Abma and Widdershoven (2008) argue that “*evaluation is not only a scientific and technical affair, but a social practice in itself*” (p.121); whereas for Saunders *et al.*, (2011),

evaluation is characteristic of all social policy areas; involves dimensions of evaluative practice consisting of symbolic structures, particular orders of meaning in particular places and has unintended effects. It consists of practices

which use implicit, tacit or unconscious knowledge as well as explicit knowledge; **can have progressive enabling characteristics but are also perceived as part of the “surveillance culture”**. (Saunders *et al.*, 2011, p.4, **emphasis added**)

The standpoint taken in this thesis is that evaluation is a SP, because it is “*undertaken by people, within structures of power and resource allocation*” (ibid. p.3). Thus practitioners’ experiences of evaluation are seen as a routinised daily practice – an evaluation practice (Reckwitz, 2002; Saunders *et al.*, 2011; Shove *et al.*, 2012).

2.3.2 Practice

A primary concern of most theories of evaluation is its practice (Christie, 2003; Saunders *et al.*, 2005; Rogers, 2008; Carman and Fredericks, 2010; Patton, 2010), but different theoretical traditions locate the social in different units of analysis. A SP perspective locates the social in the ‘practice’, which means that what one does, thinks, knows and wants on a routinised basis, is evaluative in nature - an evaluation practice. In the present research, the term, ‘practice’ is defined in terms of Reckwitz’s (2002) SPT stance, as follows:

A ‘practice’ is a routinized type of behavior which consists of several elements, interconnected to one other: forms of bodily activities, forms of mental activities, ‘things’ and their use, a background knowledge in the form of understanding, know-how, states of emotion and motivational knowledge. **A practice – a way of cooking, of consuming, of working, of investigating, of taking care of oneself or of others, etc.** – forms so to speak a ‘block’ whose existence necessarily depends on the existence and specific interconnectedness of these elements, and which cannot be reduced to any one of these single elements. (Reckwitz, 2002, p.250, **emphasis added**)

Within this thesis it carries a social perspective stamp, as a “*set or cluster of behaviours forming ways of thinking and doing associated with undertaking evaluative activity*” (Saunders *et al.*, 2011, p.2). The SP branch of the theories of practice (discussed in chapter three) focuses on the way a practice itself, in whatever domains, become an object of

examination; so what people do on a daily basis is what is termed by *practice*; and all societal life can be interpreted as sets or clusters of practices within different contexts.

2.3.3 Evaluation Practice

Scriven's (1991) idea that evaluation is a "*new discipline, but an ancient practice*" (p.5) challenged the separate use of both terms, 'evaluation' and 'practice'. It is possible that the combination of the new and ancient elements represent his understanding of EP, but there are no certainties. What is known is that a multitude of views prompted scholars to clarify exactly what they mean by 'evaluation' and by 'practice' (Saunders *et al.*, 2011). Following on from Scriven's (1991) theorisation, other scholars have provided their views on how fundamental EP is, as a property of the conceptualisation of evaluation as a discipline (Patton, 2002; Rossi *et al.*, 2005; Morra-Imas and Rist, 2009), for it constitutes the essential aim of the evaluation theories. Also, Shadish (1998) has put forward the contingency theories of EP to emphasise that:

All approaches to evaluation involve trade-offs among the many goals we try to maximize in evaluation (e.g., the goals of use, constructing valid knowledge, valuing, assisting in social change, etc.), **so we need to have conceptual tools to help us understand those trade-offs. The contingencies in which we are interested in evaluation theory are those with implications that make a difference to evaluation practice.** (Shadish, 1998, p.8, **emphasis added**)

Another core contingent to EP relates to the attribution of value to what happens in a specific time and in a given context (Mark and Shotland, 1985; Morra-Imas and Rist, 2009; Scriven, 2012; Freeman and Hall, 2012), alongside the attribution of worth or merit, as important for the society to define what is worthwhile – "*how its agencies attribute value to its policy and program interventions, how institutions decide on the quality and merit of its internal practices and how groups of stakeholders decide on the value of what they are doing*"

(Saunders *et al.*, 2011, p.1). In this thesis, the term ‘evaluation practice’ means a “*social practice bounded by the purpose, intention or function of attributing value or worth to individual, group, institutional or sectoral activity*” (ibid. p.3), so the notion of ‘worth’ in evaluation is relevant for this research, because it challenges how evaluation is perceived, mainly by practitioners.

2.4 Mapping the landscape of evaluation practice: theoretical perspectives

This section outlines the existing literature on EP, particularly focusing on its theoretical views and its approaches to research. It identifies the main theoretical views of evaluation by mapping previous empirical studies representing various approaches to the research on EP.

The relationship between evaluation theory and EP has been amongst the most crucial debates in the discipline of evaluation (Leviton and Hughes, 1981; Shadish and Epstein, 1987; Shadish, 1998) and subsequently there has been increased consensus about the need for further research (Christie, 2003; Donaldson and Scriven, 2003; Henry and Mark, 2003; Saunders, 2006). Even in 1980, it was said that “*evaluation-more than any science-is what people say it is; and people currently are saying it is many different things*” (Glass and Ellett, 1980, p.121).

By building upon these works, I identify that this link between evaluation theory and practice is likely to shape the future of EP in NPOs, because it may clarify that individuals’ activities, in one way or another, have some theoretical reasoning, that hopefully has been previously applied in their EP (even if they are not aware of it). Everybody holds an ontological position, but some might not be aware of that influence in their everyday practice of evaluation.

From 2006 onwards, recent studies have attempted to understand the initial concepts upon which researchers form their evaluation knowledge base (Chelimsky, 2013; Leviton, 2015);

however these attempts required further insights into the intersection between evaluation theory and practice. Despite early theorists' efforts, it has been said that today, evaluation theory and practice "*interact insufficiently*" (Chelimsky, 2013, p.91); nonetheless, previous works of some evaluation theorists have emphasised important contributions in this emerging research domain - linking evaluation theory with evaluation practice - as illustrated in Table 2.1 (overleaf).

Chronology	Theoretical perspective	Theorists	Main Topics
1970s	Program evaluation	Wholey <i>et al.</i> ([1970] 2010)	Funders' evaluation practices
	Values-pluralism	MacDonald (1978)	Democratic and social justice view of evaluation
	Evaluation utilisation	Patton (1978); Alkin <i>et al.</i> (1979)	Use and usability of evaluation results;
1980s	Social practice	Stake (1983)	Responsive to stakeholders' concerns
	Evaluation utilisation	Alkin and Ellet (1985)	Approaches to evaluation practice
	Constructivism; Values-pluralism	Guba and Lincoln (1988); Eisner (1985); Simons (1987)	Humanisation of evaluation; naturalistic enquiry; participatory approach to evaluation; democratic and social justice approach to educational evaluation practice
	Evaluation utilisation	Cousins and Leithwood (1986); Shadish and Epstein (1987)	Empirical research on the use of evaluation; utilisation of evaluation; evaluators examining their own evaluation practice
	Program evaluation	Greene (1988)	Stakeholders' participation; democratisation of evaluation
1990s	Program evaluation	Alkin and House (1992); Greene (1999)	Evaluation of programs; limitations of performance measurement approach; participatory approach to evaluation
	Values-pluralism	Smith (1993); Simons (1996)	Lessons learnt from evaluation practice
	Realism	Cousins and Earl (1995) Julnes and Mark (1998)	Links evaluation theory and practice; evaluation as sense-making
	Learning	Lave and Wenger (1991) Wenger (1998)	Full engagement and "identity of participation" (Wenger, 1999) in evaluation
2000s	Realism	Mark <i>et al.</i> (2000)	Assisted evaluation
	Learning	Bonamy <i>et al.</i> (2004) (Saunders <i>et al.</i> , 2005)	Provisional stabilities in evaluation practice
	Values-pluralism	Abma (2000); Mertens (2005); Simons (2003, 2009) Berner and Bronson (2005) House (2013)	Conflicting stakeholders' relationships in evaluation; social justice view of evaluation; inclusivity as the focus of evaluation practice, transformative evaluation; democratic evaluation; evaluation and/as social relation; participatory approach to evaluation

2000s (cont)	Evaluation utilisation	Henry and Mark (2003); Taut and Brauns (2003); Mayne (2008, 2010); Christie and Fleischer (2010)	An agenda for research on evaluation; resistance to evaluation; contribution analysis to evaluation; design and methods used in evaluation studies
	Program evaluation	Christie (2003)	Evaluation practice mapping onto evaluation theory
	Social practice	Saunders (2000); Becher and Trowler (2001); Stake (2004); Saunders <i>et al.</i> (2005)	Reflexive planning tool for evaluation (RUFDATA); Evaluation as a bridging tool; academic tribes and the role of disciplines
	Complexity	Stame (2004); Rogers (2008)	theory-based evaluation and types of complexity; program theory to evaluate complicated and complex interventions
	2010s	Learning	McCluskey (2011)
	Social practice	Saunders <i>et al.</i> (2011); McCluskey (2011); Trowler <i>et al.</i> (2012); Trowler (2013); Chelimsky (2013)	Reconceptualisation of evaluation in HE; evaluation moments; academic tribes and territories; practice-based ethnographies; evaluation theory and practice in the real world
	Values-pluralism	Parry-Crooke (2014); Simons (2015)	Constructivism; social justice view of evaluation; ethics of evaluation; equity-driven evaluation
	Complexity	Patton (2010); Bamberger and Segone (2011); Freeman and Hall (2012); Westhorp (2013); Mowles (2014); Piccioto (2015)	Developmental evaluation; equity-focused evaluation; complexity in a responsive evaluation; complexity in a realistic evaluation; complexity sciences in evaluation scholarship; global trends in evaluation

Table 2.1: Chronological list of the main studies linking evaluation theory and practice

Table 2.1 maps out some of the evaluation theorists' original work and the topics commonly associated within the debate of evaluation theory and its connections with EP. It represents a sample from an extensive list of published studies linking evaluation theory with EP; alongside the focus on their theoretical views. One note regarding the chronological scope is that, despite the contemporaneity of the debate, its main foundations report back to the 1970s; so for this reason I purposively extended the period covered for my review of the literature (section 2.2).

From the theoretical perspectives mapped above, the strands of program evaluation, values-pluralism, learning, and social practice appeared relevant to my research because these will frame the research on EP and their current debates presented in the thesis, as explained in section 2.5. Other specific strands of realism, evaluation utilisation and complexity were not fully explored because, despite their relevance to evaluation, their theoretical perspectives have not been focused, exclusively, on individuals' practice of evaluation. This process of narrowing the review has inevitably shaped the boundaries of this research.

The summary of the main theoretical perspectives employed to examine EP is visualised in Table 2.2 (overleaf) and sets the review, emphasising how each perspective has been previously used, its limitations, what I take from each and why.

Theoretical perspective	Original work	Approaches to evaluation further developed by:
Evaluation utilisation	Patton (1978); Alkin <i>et al.</i> (1979); Alkin and Ellet (1985)	Cousins and Leithwood (1986); empirical research on the use of evaluation Shadish and Epstein (1987); evaluators examining their own evaluation practice Henry and Mark (2003); an agenda for research on evaluation Taut and Brauns, (2003); resistance to evaluation Mayne (2008; 2010); contribution analysis to evaluation Christie and Fleischer (2010); design and methods used in evaluation studies
Program theory	evaluation Wholey <i>et al.</i> , (2010); Scriven (1991); Rossi <i>et al.</i> (2005)	Patton (1997); utilised-focused evaluation Greene (1999); inadequacy of performance-based approach to evaluate NPOs Carman (2007); Rogers <i>et al.</i> (2000); Mayne (2001); performance-based evaluation Rogers (2008); program theory to evaluate complexity Morra-Imas and Rist (2009); results-based oriented evaluation
Values-pluralism	Guba and Lincoln (1987;1989)	Fetterman (1994); empowerment evaluation Greene (1988); Abma (2000); stakeholder evaluation; participatory evaluation Mertens (2001; 2005); transformative evaluation Simons (1987; 2003); democratic evaluation Stake (1995; 2004); responsive evaluation Kushner (2000); personalising evaluation Bamberger and Segone (2011); equity-based evaluation
Realism	Pawson and Tilley (1997)	Pawson and Manzano-Santaella (2012); realistic evaluation

		Julnes and Mark (1998); Mark <i>et al.</i> (2000); evaluation as sense-making
Learning		Bonamy <i>et al.</i> (2004); metaphor of the 'implementation stair case' in EP
	Lave and Wenger (1991);	McCluskey (2011); evaluation moments
	Wenger (1998)	Trowler (2013); practice-focused ethnographies
		Schwandt (2015); a life for the mind of practice in evaluation
Social practice		Abma and Widdershoven (2008); evaluation as social relation
		Saunders (2000); practical approach to evaluation planning
	Giddens (1979); Schatzki	Saunders <i>et al.</i> (2005; 2011); evaluation creating provisional stabilities; reconceptualising evaluation in HE
	(1996); Schatzki <i>et al.</i>	McCluskey (2011); evaluation moments in a practice-based evaluation
	(2001); Reckwitz (2002)	Trowler <i>et al.</i> (2012); academic tribes and territories
		Trowler (2013); practice-based ethnographies
		Chelimsky (2013) knowledge-based evaluation
Complexity		Patton (2010); developmental evaluation
		Bamberger and Segone (2011); equity-focused evaluation
	Patton (1994); Stame	Freeman and Hall (2012); complexity in a responsive evaluation
	(2004); Rogers (2008)	Westhorp (2013); complexity in a realistic evaluation
		Mowles (2014); complexity sciences in evaluation scholarship
		Piccio (2015); democratic evaluation in complex times

Table 2.2: Main theoretical perspectives to research evaluation practice

Drawing on Table 2.2 I acknowledged the four relevant theoretical strands to my research: program evaluation theory, values-pluralism, learning and social practice. As these have been previously reviewed elsewhere (Shadish *et al.*, 2000; Guba *et al.*, 2011), rather than repeat it, I will emphasise why they are important to this study and their potential limitations.

Program evaluation theory (PET) is important as it is the current theory applied by funders to evaluate NPOs' social interventions (Carman, 2007). As this research is an exploration of a DE publicly funded project, it matters that a PET advocates for a performance-based approach to evaluation (Rogers *et al.*, 2000; Carman, 2007, 2009), results-oriented (Morra-Imas and Rist, 2009) and grounded in attribution analysis (Mayne, 2001), (see Appendix 3; glossary). Thus, a critique of the PET relates to the social scope of the NPOs' interventions with commentators arguing that the performance-based approach to evaluate their social interventions is inadequate (Greene, 1999; Stevenson *et al.*, 2002). Preliminary work has been qualitatively undertaken by Greene (1999), who argues that the constructed idea of accountability is regressive and there is need for revitalising these ideas of results-oriented and program effectiveness. As a social constructivist scholar, and Cronbach's (1980) former student, she has been influenced by his motto that "*a demand for [exclusive] accountability is a sign of social pathology in the political system*" (1980, p.4 **[emphasis added]**). In this vein, Greene (1999) critiques the performance approach as a distortion to respond to the accountability demands of the 20th century. So these systems, as applied to evaluation, sometimes fail to capture or represent program quality, as well as ignore important features such as, the participants' experiences, the context, and the project's provisional contribution in achieving its funded

outcomes (Saunders *et al.*, 2005; Mayne, 2010). These features appear in a NPOs context of evaluation, as further discussed in section 2.5.2; NPOs and performance.

Evaluation theorists advocate that evaluation is understood through different approaches, thus it is practised differently across contexts, disciplines and circumstances (Greene, 1999; Henry and Mark, 2003; Donaldson and Scriven, 2003; Shaw *et al.*, 2006; Abma and Widdershoven, 2008; Vo and Christie, 2015). These differences may relate to distinct theoretical standpoints, diverse motivations to evaluate, and opposing conceptions of the purpose of evaluation. For instance, a values- pluralism approach (VPA) accepts multiple realities of the world, and in this sense, offers an alternative to the previous critique of a PET performance-driven. A VPA enables the emergence of several attempts to transform the practice of evaluation into a more just and equitable practice; in other words to humanise evaluation (Lincoln and Guba, 1985; Abma, 2000; Stevenson *et al.*, 2002; Simons *et al.*, 2003; Mertens, 2005; Abma and Widdershoven, 2008), particularly, regarding the shift of focus from a PET towards participation of people within the evaluation processes – a democratic participation (Simons, 1987) (see section 2.5.3.1; participation of stakeholders). As illustrated in Table 2.2 above, some theoretical foundations of VPA and learning draws on the participation of stakeholders (Greene, 1988; Abma, 2000), the idea of democratic evaluation (Simons, 1987; 2003) and its personalisation (Kushner, 2000). These ideas are critical to my exploration of EP, because they support and advance an approach of evaluation that draws attention “*to the program uniqueness and to the social plurality of its people*” (Stake, 2004, p.8) – a social practice approach to evaluation - that has been developed by Saunders,

Trowler, and Bamber (2011), drawing on Stake's (ibid.) responsive evaluation. They reconceptualised evaluation as a social practice in a setting where multiple stakeholders operate and this responsive approach has its origins in a participatory view of evaluation, particularly, regarding the stakeholders' voices being heard, as well as their engagement throughout the evaluation trajectory (Saunders *et al.*, 2011). They were also informed by the theories of learning in evaluation, namely through Wenger's (2008) communities of practice. The main point taken from this learning approach is that the more practitioners engage and participate in evaluation, the more they learn from it. The sharing of their experiences prompts the circulation of knowledge and their identities are likely to be reinforced (ibid.). Two other notions of theories of learning - practitioners as carriers of evaluation and their full engagement in evaluation - are further developed in chapter three, about theories of practice.

In this vein, a social practice theoretical (SPT) view of evaluation appears critically important to my research, because it holds the space and opportunity for stakeholders (and practitioners) to influence, and ultimately change, the evaluation route initially planned, through their verbalised opinion (Saunders *et al.*, 2011). This view is relevant in a context where an increased request for accountability has shaped current structures of power and held divergent perceptions and expectations about participation and EP, as it is the context of NPOs (Kearns, 1996; Carman, 2009). However, a limitation of Stake's (2004) responsive approach lies in the assumption that a practitioner is given the opportunity to participate, give feedback and have their voice heard during evaluation processes; which in many NPOs of the social sector, is not yet part of their main tasks (Carman, 2007). Some evaluation theorists have critiqued this

apparent responsive evaluation, by arguing that the key point of participation in EP is to empower participants; and if this have not occurred evaluation is making a disservice to improve societal life (Schwarz and Struhkamp, 2007; Flores, 2008; Schwandt, 2015).

In summary, in the past two decades, a number of researchers have sought to determine alternative theoretical approaches to evaluation (Greene, 1999; Mertens, 2001; Fitzpatrick *et al.*, 2004; Saunders *et al.*, 2005; Guba and Lincoln, 2005; McCluskey, 2011; Saunders *et al.*, 2011), because extensive literature had been published on the inadequacies of performance-based approach to evaluation (Greene, 1999; Simons, 2003; Saunders *et al.*, 2011), supported by program evaluation theory. The main critique is that its managerial process and highly outcome-driven process have failed to attribute worth (evaluate) to interventions within the social domain (Saunders, 2000; Simons *et al.*, 2003; Saunders *et al.*, 2005). The views of Saunders (2000; 2005; 2011) and Simons (2003; 2009; 2015) have informed the research study presented in this thesis and are described in detail in chapters four and five, respectively.

Whilst the above views represent a selection of perspectives to examine EP, they do not cover all the arena of theoretical positions in evaluation; for instance, other theorists have drawn on psychological views of evaluation, particularly around affective evaluation (Love, 2004) and language of evaluation (Patton, 2000). From the four perspectives selected – program evaluation theory, values pluralism approach, learning and social practice theory - I take to my research the following central ideas, to conduct an in-depth longitudinal case study on EP and how it changes over time, in a NPO domain of DE:

- Performance systems applied to evaluation fail to capture and represent the program quality, as well as ignore important features such as, the participants' experiences, the context, and the project's provisional contribution in achieving its funded outcomes (Greene, 1999).
- Stakeholders tend to assign diverse and conflicting meanings to evaluation, which generates change in the social relations within an evaluative setting (Abma, 2000).
- In many NPOs of the social sector, a non-profit practitioner is yet to be given the opportunity to participate, give feedback and have their voices heard during evaluation processes (Greene, 1988; Carman, 2007).
- The more practitioners engage and participate (in evaluation), the more they learn, in and from evaluation (Wenger, 2008).
- A learning perspective of evaluation brings the human interface to the front of EP (McCluskey, 2011).

This section highlighted that EP, as a research domain, is growing, in order to accommodate developments on how to evaluate in a fast-changing globalised world. To this regard, scholars have argued that there is need for further empirical research on how EP can best serve the social domain (Henry and Mark, 2003). Theoretically speaking, this review has demonstrated the call for a non-linear framework to understand how a social intervention works, how worth is attributed, and how evaluation is practised on a routinised basis (Sanderson, 2000; Stake, 2004; Perrin, 2011; McCluskey, 2011). This call has been reinforced by suggestions that a performance-based approach to evaluation appears inadequate in the social domain (Greene, 1999; Simons, 2003), which leads to an increased attention to the theories of SP and social views of learning

applied to evaluation (Lave and Wenger, 1991; Saunders, 2000; Wenger, 1998; Saunders *et al.*, 2011).

This review presents an opportunity to identify the main gaps in the literature of EP, by having mapped its theoretical origins. From an EP-research perspective, this mapping involves the current debates in the literature, to further investigate how these are developed in this thesis' research setting of DE.

2.5 Current research debates in evaluation practice

This section reviews four of the current debates in EP: first, the influence of the context in evaluation; second, NPOs and the performance-based approach to evaluation; third, the issue of power relationships within an evaluative setting and fourth, the role of practitioners within a funder-recipient based context. The connection between the previous theoretical section and the current is made by commentary on how different theoretical perspectives have been applied across a variety of contexts and circumstances.

2.5.1 The influence of context in evaluation practice

Context has long been a question of great interest in a wide variety of fields (Creswell, 2007; Wolcott, 2008; Silverman, 2011), which has led evaluation scholars to address questions concerned with its role in the evaluation (Pawson and Tilley, 1997; Christie, 2003; Saunders *et al.*, 2005; Rossi *et al.*, 2005; Simons, 2009). Despite their realistic view of evaluation, which this thesis does not subscribe to, it is important to acknowledge Pawson and Tilley's (1997) contribution on the role of context in evaluation research. They define context as *“the current background circumstances which encourage or enable a particular*

group of stakeholders to be assembled for negotiation”, (p.70) and go beyond that by specifying that by “social context” they *do not refer simply to the spatial or geographical or institutional location in to which programs are embedded*” (ibid.), and this lack of consideration to the social circumstances is one important exclusion of evaluation research. In this vein, my research will closely explore the DE social context, in which evaluation is practised.

Some scholars suggest an increased need for understanding the influence of context on the nature of EP, as Bhola (2003) remarks:

Evaluation as practice has to resonate to a **multiplicity of layered contexts**: the professional culture of evaluators and the institutional culture of the place in which evaluation is [conducted]. (Bhola, 2003, p. 391–392, **emphasis added**)

Others have focused on the particular role of changes occurring within different contexts of evaluation and how these changes impact upon EP (Smith, 2009). For instance, “*it is important to understand the characteristics of these contexts that are now beginning to intersect*” (ibid.p.44). The influence of context has also been understood as a major force that has shaped evaluation around the world (Westley *et al.*, 2006; Patton, 2010), in which the role of governmental policies and practices has been assumed as the most influential contextual force, leading to influence multiple forms, approaches and methods of evaluation (Brown *et al.*, 1985; King, 2007). As an example, Karlsson (2003) notes that in the European context, intergovernmental forces such as the European Union have profoundly shaped evaluation. In this respect, Smith (2008) argues that:

when the history and influences of these [global] forces are disparate across settings, increased spread and interaction are likely to result in transformation. (Smith, 2008, p. 44)

Despite the useful mapping of the fundamental issues for evaluation in a globalised world, Smith (ibid.) fails to address practical recommendations to evaluators and practitioners within a global society. The study acknowledges many challenges that evaluators and practitioners might face; recognises that context matters; and subtle changes are likely to appear, changing the fundamental nature of evaluation, both in individual contexts and worldwide.

Although past scholars have contributed to the debate, still, there is much to be learned about context, and scholarly work remains at the early stages of this process (Vo and Christie, 2015). Vo's (2013) comparison on how context is perceived in three areas of evaluation theory – participatory, values-engaged, and realist – suggests that issues related to the organisational aspect of context are of great importance in participatory approaches, but are deemphasised in a values-engaged approach. Later, Vo and Christie (2015) describe the evaluation context as a dynamic and multi-layered phenomenon, and notes that, possibly, because context includes many issues, few comprehensive studies have looked at it in-depth.

2.5.1.1 Funder-recipient based context

Relevant literature has noted that the role of funders and how they request evaluation is a major area of interest within the field of EP (Stevenson *et al.*, 2002; Newcomer *et al.*, 2004; Carman, 2009; Carman and Fredericks, 2010). In this regard, there have been a number of studies referring to funder-recipients, stakeholders or grant-holders (Greene, 1988; Morra-Imas and Rist, 2009; Saunders *et al.*, 2011; Berry *et al.*, 2015), in the context of NPOs (Carman, 2007) and in other institutions (Saunders *et al.*, 2011). The term, “*funder-recipient based context*” was coined by Saunders, Trowler, and Bamber within

the evaluation context of HE (2011, p.12), and for the purpose of the thesis, it represents the funding agency and the organisation in receipt of funding, respectively. The relational nature of evaluation dominates the type of relationships within a funder-recipient based context, because evaluation is frequently set against a political landscape, with power relations included (Simons, 1987) as an intrinsically political (Shadish *et al.*, 2000) practice. As such, the previous debate around NPOs and their performance intersects with this debate of power, calling for a review of DE as a funder-recipient based context (examined in section 2.5.21; DE NPOs).

NPOs social interventions can be designed and implemented in preparation for a project, program or policy scheme; which may vary in their durability and depend on funders' guidelines. However, there are situations in which interventions with different timescales are termed projects, and funded under specific schemes, because most external agencies have project cycle guidelines set to conceptualise these schemes. For instance, the Department for Co-operation of the European Commission (EuropeAID) defines 'project' by mentioning not only its domain, but the implementer, the local partners, the financial cost and the timeline (EuropeAID, 2004). The Department for International Development and Cooperation (DfID)⁵, in the UK, tends to request a *logframe* or logical framework of the intervention, which is a matrix with its goals, outcomes, outputs, indicators and assumptions (see Appendix 2b; example of a logic framework). Usually, NPOs design a *logframe* when formulating the intervention to indicate their long-term goals, the changes to be achieved and

⁵ See www.dfid.co.uk

how, by when, and with whom they have to be implemented. To do so, the software, Microsoft Project helps to manage interventions across project timelines, supporting the effective attribution of human and financial resources to each activity. This cycle of a project is termed project cycle management (PCM). A *log frame* as a feature of a funder-recipient based context is required by funders from the start of the project, so it is likely to require the involvement and knowledge of practitioners. Whilst, the literature assumes that practitioners in a social domain are able to fill in *log frames* and manage projects from a PCM, performance-based approach (Carman, 2007), it has been demonstrated that the implementation of interventions in a funder-recipient based context requires a sequence of stages (EuropeAID, 2004), as illustrated in Table 2.3, below:

Stage of the cycle of the project	What is it?
Identification	Includes needs assessment
Formulation	Includes project design
Implementation	Includes monitoring and management
<u>Evaluation</u>	Includes monitoring and reporting
Audit	Includes regulation and inspection

Table 2.3: Stages of the cycle of the project (EuropeAID, 2004, p.33)

There are a series of stages that, if funding is granted, have to be followed according to funders' guidelines (EuropeAID, 2004), and the information generated across these stages has been known as "*performance measurement*

information” (Carman, 2007, p.71). Despite minor differences, the majority of funders follow the stages of the project cycle as the figure above. The debate around a performance approach to evaluation has been whether, and how, NPOs use this approach in their EP; thus for the purpose of this study, I further explore the PCM approach used by funders – a performance approach – to define the power relationships at stake in the funding-recipient context of this research.

2.5.2 Non-profits organisations and performance

In order to understand how organisations evaluate their interventions, it is relevant to first define the organisations that evaluate, so this study focuses on the exploration of NPOs, recognising, however, that many profit-making organisations also evaluate their interventions. NPOs are defined as a-political organisations, not making or conducted primarily in making a profit (Simpson *et al.*, 2009) and, in the UK, are also known as charities or charitable organisations (MacDonald, 2011). These are defined as organisations set up to provide help and raise money for those in need (Simpson *et al.*, 2009).

Extensive literature has been published on the development of NPOs, their structure, management, and role in the society (Brody and Cordes, 1999; Salamon, 1999; Jeavons, 2005; Carman, 2009; Murray, 2010; Carman and Fredericks, 2010; Thomson, 2010) ⁶. For the purpose of this thesis, NPOs operate in a context in which a funding grant has been successfully acquired from an external agency and a subsequent evaluation is demanded. With some

⁶ For a further review of literature on NPOs, see LeRoux and Wright (2010) and Lee and Nowell (2015).

variations, a NPO's mission tends to contribute to the world's greater good, either through health projects, social cohesion programs or even through raising awareness about the global world and its interdependencies. Whilst some NPOs are recognised as activists, others are doers; some prioritise training and supporting communities, while others work primarily with children (MacDonald, 2011). Many other forms of philanthropy have emerged and sometimes NPOs are also considered foundations; however, an interesting feature is that, in the US, two thirds of foundations have staff specifically assigned to evaluation (Carman and Fredericks, 2008). What is known about NPOs EP come primarily from a large scale project - Voluntary Sector Evaluation Research Project – in which a telephone survey was conducted with 1,965 NPOs and 322 funders in 2001, which found that the evaluation expectations amongst funders had increased between 2005-2008, “*yet less than half of the funders provided funding for evaluation activities*” (ibid, p.53). Arguably, the core mission of NPOs seemed to dictate the design of their interventions, the funding requested, the practitioners employed, and the type and purpose of the evaluation. This review narrowed the scope to a specific type of NPOs – Development Education (DE) NPOs, because it is the context in which the phenomenon of EP is explored.

2.5.2.1 Development education non-profit organisations

The importance of EP as an emergent research domain, is unquestionable, thus NPOs were brought into this debate in the last decade (Carman, 2007; Carman and Fredericks, 2008). Besides, in an uncertain world, what is taken-for-granted today, will be challenged tomorrow, if not before, thus the literature has emphasised how NPOs' practice of evaluation is shaped by contemporary

challenges. The practice of evaluation aims to improve the quality of interventions delivered, through learning the strengths and weaknesses of each particular case.

DE is a recent discipline in the interdependent and globalised world, aiming to raise awareness and understanding of how global issues affect the everyday lives of individuals' communities and how each one of us can and do influence the global (DARE Forum, 2004). (See Appendix 2a; background information on DE). As no common understanding of a DE standard definition, mission, and values has been reached, a concern about its shared vision has increased among its practitioners (Bourn, 2011). Consequently, the practice of evaluation appears particularly important, because NPOs are more likely to encounter difficulties in getting reliable evidence of their actions (in aiming to develop awareness, to increase understanding, and to ensure a commitment that leads to attitudinal change). Thus, an adequate evaluation approach is vital in ensuring that the attribution of worth in DE interventions is correctly made (Bourn, 2014). DE organisations tend to promote education for a sustainable world, and one of their main distinctive characteristics is to ensure that the civil societies of most developed countries are aware of the reality of the least developing countries⁷, in terms of poverty, food security, education, health and environment. Once the developed countries are aware of the reality of the least developing countries, their citizens are expected to act in order to promote justice, fairness and ultimately a social change of attitudes, making the world a better place. However, these characteristics of DE are problematic when it is time to evaluate

⁷ See <http://hdr.undp.org/en/reports/global/hdr2010/summary/>

their social interventions, because organisations will have to attribute value to whether and how social change has been achieved.

2.5.2.2 Development education evaluation practice: performance-based

In recent years, there has been an increased interest in DE in general (Coriddi, 2008; Courtney, 2008; Holland and Thirkell, 2009; Krause, 2010; O'Brien, 2011; Bourn, 2011, 2014); however, the sector's tradition of EP seems fragile and still performance-based, as funders require hard evidence to demonstrate the achievement of results (Bourn, 2011; 2014). Literature suggests that DE practitioners struggle to evaluate in the current sector's tradition of performance, because the nature of the interventions evaluated are socially unpredictable, and complex; therefore the performance information does not seem to capture the essence of DE interventions (Bourn, 2014). Consequently, DE scholars have raised questions about whether to maintain a well-established performance-based evaluation approach or to apply an alternative approach to evaluation. Nevertheless, literature regarding how evaluation has been practised in the DE domain is slim; but suffice to say that it is a fragmented discipline in need of further research (Bourn, 2008; Coriddi, 2008; Bourn, 2014), particularly, to *“place practice at the heart of the evaluation in DE”* (Bourn, 2014, p.50).. For instance, over the past decade, recent changes have been experienced, regarding funders' requests for evaluation. DE organisations in receipt of external funding have been requested to evaluate using the conventional funders' approach of performance-based evaluation, widely used in a variety of sectors (health; education and HE).

Whilst some have been collecting data to illustrate performance-based indicators (Carman, 2007), others, in the DE sector, appear to have not done, or to struggle to do so (Bourn, 2011). Although contemporary approaches to evaluation such as, transformative, equity-based and democratic evaluation, have recognised new nuances of competence in terms of arts-based methods (e.g. (Simons, 2009), these do not yet appear to be established in a DE context (Bourn, 2014). The gap emerging from this debate is the lack of qualitative exploratory studies examining how DE organisations respond to a performance-based evaluation request.

Furthermore, in looking at funders' performance-based approach in these sectors, past commentators have raised questions about its appropriateness as the standard approach to evaluation (Guba and Lincoln, 1988; Greene, 1999). This question has been, and still is, debated as especially relevant, because in domains like DE, the purpose of evaluation is to show evidence of attitudinal and social change achieved. Another issue emerging from the slim literature on DE is whether the performance-based approach to evaluation seems to do a disservice to its EP. The answer is twofold: first, the rationale to demonstrate impact through a performance-based approach leads to NPOs accountability of public funding received (Carman, 2007); second, the particularities of the social domain, in terms of demonstrating evidence, conflicts with the performance approach, because DE's mission is to achieve social change in the long run (Greene, 1999; Simons, 2003).

The need for further research on EP in the DE domain is evident in the analysis of three works: first the European Development Education Monitoring report

(Krause, 2010); second, Kaimacuata's (2010) paper on challenges implementing DE projects for excluded students of primary and secondary education; and third, recent work from Douglas Bourn (2008; 2011; 2014), a notable exception of high-standard academic work in the DE domain.

From these readings, it was highlighted that an implication of the lack of conceptual consensus about DE, was the UK Government's move to reduce DE projects in formal education⁸. As a result, the Coalition government in 2010 adopted a new approach to DE evaluation, arguing that the link between these DE interventions and poverty reduction was not strong enough to satisfy their rigorous criteria for development impact (O'Brien, 2011). Another implication was that this DE shift had consequences on its funding, because "*development education has moved from an expansionary to a contractionary or survivalist mode*" (Khoo, 2011, p.2) in a way that general government support for DE projects was replaced by an emphasis on evaluation results and its impact (Bourn, 2011).

Consequently, some interventions were ended due to a poor or absence of demonstration of project impact or value for money. To face this survivalist landscape DE NPOs were asked to carry a double duty on the evaluation front - to demonstrate the benefit to the intended beneficiaries, while, delivering results, transparency, accountability and value for money for tax payers (Bourn, 2011).

⁸ See <https://www.gov.uk/government/announcements>, DfID, assessed on 12th May, 2015

The evaluation process of DE interventions seems particularly complex in the context of measuring outcomes, for various reasons. First, the DE context is composed of a complex set of social and educational values that necessarily requires adequate indicators to attribute the value of their outcomes (Feuerstein, 1988). Second, the discipline of DE entails the ultimate goal of development awareness, which means an increased understanding of global poverty, and a commitment to change social attitudes (DARE Forum, 2004; Adams, 2010).

Summing up, it stood out that there is a need to further investigate EP in the DE domain. It seems necessary to explore the constraints faced by DE practitioners when practicing evaluation, so there is a call for in-depth research about the contexts in which this practice operates (Kaimacuata, 2010). Finally, additional research on EP in the DE domain appears much needed to recognise this interdisciplinary domain as a credible discipline (Krause, 2010). The scant DE literature recognises the need to shift from a performance-based approach to a practice-based approach to evaluation (Bourn, 2011, 2014), which promotes the following debate on power relationships in evaluation practice.

2.5.3 Power relationships in evaluation practice

Power is a dominant feature of evaluation activity. Evaluation is perceived and experienced as “*a controlling and surveillance culture*” (Saunders *et al.*, 2011, p.4). As evaluation is a social practice, power and conflict are bound to occur. Previous studies on EP have demonstrated that one main debate is how to prioritise social relations within a power-imbalanced context (Abma, 2000; Gong and Wright, 2007; Abma and Widdershoven, 2008). Some scholars have prioritised the understanding of power relations through people’s relationships

(Atjonen, 2015); whereas others have focused on people's involvement and participation in evaluation (Wenger, 1999; Saunders, 2000; Turner, 2005). Although, a "*funder-recipient based context*" (Saunders *et al.*, 2011, p.12) represents a powerful funder and a powerless organisation, individual or grant holder; the question of how to incorporate power and conflict in a manageable way, into a practice-based evaluation tradition, remains unanswered. By a power imbalanced context, scholars mean a power inequality represented within one organisation, project, or structure (Westley *et al.* 2006), but for that representation to occur, individuals have to participate in that routinised life, by means of an organisation, project or structure.

2.5.3.1 Participation of stakeholders

As participation of stakeholders appears as a hallmark of the values-pluralism theory, it is relevant to clarify the meaning of both terms in the thesis. The meaning assigned to participation draws on Wenger's (1998) theory of communities of practice (CoP), to emphasise the feature that participation both shapes individuals' trajectories and involves context or communities being shaped by these individuals, as they engage in practice. Their full participation is implicitly assumed when Wenger theorises about the process of learning acquisition. Lave and Wenger (1991). have earlier noted that "*the production, transformation, and change in the identities of persons, knowledgeable skill in practice, and communities of practice are realised in the lived-in world of engagement of everyday practice*" (p.47).

Stakeholders are "*people whose lives are affected by the program and people whose decisions can affect the future of the program*" (Gold, 1981; 1983 cited in

Greene, 1988). Within these groups various profiles are considered such as: program managers; funder workers; program practitioners; board members and evaluation sponsors, as well as other groups left out from participation. This broad definition of stakeholders takes into account the need to be specific about which groups of people are considered within this label, because different evaluative endeavours - funding agencies, volunteers or practitioners – can be considered as stakeholders. Equally, these groups tend to have opposing, sometimes, contradictory perceptions of evaluation, as its practice is highly-driven by the social constitution of the group. According to Stake (1983) stakeholders are viewed not as specialists or advisers, but as “*active, engaged collaborators*” (p.18). These notions around the definition of stakeholders, alongside engagement and full participation, are relevant to explore participants’ experiences of evaluation in this research.

The participation and collaboration of stakeholders in evaluation has been, and still is, a major area of interest within the field of EP (Ayers, 1987; Cousins and Earl, 1995; Simons, 1987; Greene, 1988; Mark and Shotland, 1985; Mertens, 2005; Simons and McCormack, 2007; Simons *et al.*, 2003), and collectively, these studies outline a critical role for stakeholders’ participation in evaluation. During the 2015 International Year of Evaluation⁹, this role has been reinforced by the increased attention given to democratic forms of practising evaluation.

⁹ See <http://unhabitat.org/2015-international-year-of-evaluation/> assessed on 12th February 2016.

2.5.3.2 Leadership in the funder-recipient based context

Leaders are critical players within unpredictable contexts of EP, especially when the evaluation demand is established against a political background (Saunders *et al.*, 2011). However, in a funder-recipient based context, the role of the leader is manifested in two dimensions: first, as a leader of a NPO or as a manager (Kunreuther *et al.*, 2008; Renz and Herman, 2010), and second, in the capacity of funder, as a leader within the context in which funding policies operate (Hoole and Patterson, 2008; Carman and Fredericks, 2010).

Regarding the first dimension of a leader of a NPO, the literature on leadership for non-profit management is extensive and has been widely recognised in the evaluation domain (Collins, 2005; Jeavons, 2005; Kunreuther *et al.*, 2008; Alaimo, 2008; Renz and Herman, 2010). Previous studies focused on how NPOs, as social change organisations should adapt their leadership style to a fast-changing world (Kunreuther *et al.*, 2008), providing these organisations are committed to creating a just and sustainable world, to work with communities, to promote the youth, to advocate for policies, and to raise awareness with a view towards reducing injustice (Fullan, 2011). Scholars who have attempted to define the future of non-profit leadership advocate that the subjects faced by social change organisations tend to be similar to those faced by NPOs. Specifically, Kunreuther *et al.* (2008) mentions some internal and external influences in the career ladder of non-profit leaders:

The reluctance of many younger nonprofit practitioners to assume executive director positions — to take their place in the pipeline when opportunities emerge — has much to do with all the demands of fundraising, practitioners management, board maintenance, and external relationships.(Kunreuther *et al.*, 2008, p.9)

The element of power implicitly present in the relationships of leaders is also linked with younger practitioners' reluctance in taking on leadership roles and this may be explained by their desire to achieve social change. Nowadays, a non-profit leader has to be a "*change leader*" (Fullan, 2011, p.23) one that learns by experience, and has the willingness to reflect on the evaluation practised due to their own motivation and not because it has been imposed by the funders; on the contrary, those impositions are said to shrink leaders' "*capacity, innovation, and creativity*" (ibid.). Successful leadership has been closely associated with practice.

The advice for change leaders that we have established thus far is to dwell on **your own situation and practice**—as well as that of other practitioners—as a basis for action. (Fullan, 2011, p. 27 **emphasis added**)

Fullan's position is that each leader has to prioritise their everyday practice and develop their career ladder from a practice-based perspective. Other commentators suggest that it is the commitment of the leader that matters; not only the context and the theoretical approach to evaluation, as suggested below:

the commitment of organisational leadership is critical in transforming the role of evaluation from one of basic reporting and accountability to a true process of continuous organisational learning. (Hoole and Patterson, 2008, p.93)

Although Hoole and Patterson (2008) acknowledges the critical role of leadership in organisational learning, there is a consensus among commentators that leadership is especially important in times of change and crisis (Greene, 1988; Westley *et al.*, 2006; García-Iriarte *et al.*, 2011) and that the structures of power are likely to change accordingly, influencing the social relations developed (Abma, 2000; Abma and Widdershoven, 2008).

In an exploratory qualitative study about the social relations amongst stakeholders in evaluation, Abma and Widdershoven (2008) advocate the centrality of social relations in leaders' EP:

A leader influences what people think and do, inspires and motivates, reduces ambiguity and uncertainty and resolves conflicts. If leaders lack credibility, their impact on the programme can be negative. (Abma and Widdershoven, 2008, p.213)

Leadership, as a social relation, influences practitioners' relationships; however these social relations operate in a context of an imbalance of power, which may lead to novel challenges for EP, as noted by Mintzberg (1980):

it is time to recognize that **managing is neither a science nor a profession; it is a practice, learned primarily through experience, and rooted in context.** (Mintzberg, p. 9, **emphasis added**)

The practice of managing an organisation is also the NPOs leaders' role, according to Alaimo's (2008) study, which examined non-profit leaders' previous knowledge and perception of evaluation; these features are summarised in Table 2.4 (overleaf):

Features of non-profit leadership (Alaimo, 2008)	Brief description
Hierarchical position	Typically the highest-paid staff in a non-profit organisation
Leaders driving the process	Leader plays an important role in shaping the organisation's vision
Managing the political environment	Challenge of being accountable to multiple stakeholders - some of them with competing demands for the leader and their organisation – to which leader balances these responses while managing the strategic direction for the future
Incorporating organisation learning	Influencing staff to achieve a common goal, remain competitive, and reach their full potential
Decision-making and change	Leaders must be able to handle the personal and professional discomfort that typically comes with the reflection and introspection necessary for effective cultural change

Table 2.4: Features of non-profit leadership (source: Alaimo, 2008)

In Alaimo's (2008) analysis, these features appear to shape leaders' EP, as well as influence staff's attitudinal response to EP. For the purpose of my research, these features are critical to exploring how a DE leader operates and how they influence (or not) their (and other's) EP over time.

Regarding the perspective of funding leaders, the literature on this topic has been populated by a myriad of relevant issues such as: the absence of evaluation culture in recipients' organisations (Greene, 1988); the (in)compatible expectations between funders and recipients (Henry and Mark, 2003); the increased demand of performance measurement information (Hoole and Patterson, 2008); the nature and type of funding (Carman, 2007, 2009), and the funders' decision making process on grant management (Wholey *et al.*, 2010). Besides, previous studies mentioned that the importance of culture has been

underestimated in how organisations function and that “*culture needs to be observed, more than measured*” (Schein, 1996, p.229).

The discussion around the role of the funder and their demand for evaluation has been challenged by Carman’s survey (2007) on how NPOs respond to the increased funders’ demand for evaluation. The survey’s findings demonstrate that many NPOs still struggle with funders’ requirements, particularly, the ones demanding performance information. Furthermore, organisations were not receiving additional funding to collect data for evaluation, possibly, because they had not assigned costs for evaluation activities; had not thought about evaluation, in terms of complying with regulations; or even had not conducted investigations as needed (ibid.). The findings also show that many organisations were not taking advantage of data collection strategies that could potentially help them gather useful performance measurement data. Although logic models, such as the *log frame*, may continue to receive attention in the funding and evaluation communities, they had not been used by managers at the time of Carman’s (2007) study.

The conclusion that the NPOs neither assigned evaluation costs to the funders’ budget nor used funding to improve their data collection systems, has triggered scholars’ curiosity about whether organisations communicated with funders, beyond this evaluative dimension (Henry and Mark, 2003; Carman, 2007; Carman and Fredericks, 2010). More specifically, issues like anxiety (Donaldson *et al.*, 2002), low-trust (Schwarz and Struhkamp, 2007; Dane *et al.*, 2012), and fear (House, 2013) have been described as the perceived reactions in a funder-recipient context, as well as explored in other cases of institutional evaluation

within the domain of HE (Simon and Knie, 2013). Conversely, although most investigations highlighted practitioners' reactions to the funders' evaluation demand, Stevenson *et al.*(2002) stressed the frustration of funders in recognising practitioners' limitations to respond to the evaluation demands. This aspect of the funders' frustration and the management of the expectations raised toward NPOs connect with the focus of the current research about how evaluation is practised among different groups of participants (practitioners, coordinators and funders).

To better understand the mechanisms underlying the role of funders' leadership, the organisations' reactions and their effects, some authors have investigated the link with "*compliance- based accountability*" (Carman, 2009; Carman and Fredericks, 2010). This connection focuses on the NPOs' need to conform to rules and the effects of that pressure. It also emphasises the threat of failure, reinforcing the fear of the consequences in evaluation (especially, if funders are displeased). House (2013) eloquently illustrated this threat:

The capture of evaluation by its sponsors is the greatest threat the evaluation community has faced for some time. In fact, the credibility of the field is at risk. (House, 2013, p. 64, **emphasis added**)

On the one hand, the issue of how power relationships have been manifested within a funder-recipient based context is key to the practice of evaluation (Abma, 2000); on the other hand, the decision-making process about which recipients are funded is based on many aspects, which include nuances of power such as, evaluators' influence (Kirkhart, 2000); personal relationships (Gong and Wright, 2007); and contradictions and expectations in EP (Atjonen, 2015). A

recent study that explored professional experiences of fifteen Finnish evaluators stressed the ‘expectation’ element in the work of evaluators (Atjonen, 2015), raising attention to how funders’ expectations are manifested in NPOs. Turner’s (2005) framework of how power has been exerted by people is used in Atjonen’s (2015) work to highlight the meaning of power: as the use of expertise; as interactive relations; as a fight for positions; as a reporting responsibility, and as knowledge management. The study concludes that “*evaluators’ ability to initiate, nurture, and establish relationships is crucial*” (ibid. p.44). It would have been interesting to know how power was interpreted by practitioners in a Finnish context of evaluation.

Previous studies sustained that, in line with their role as leaders, funders have been pressured from their governments to be accountable to their citizens. According to Fletcher and Dyson (2013):

we live in a political climate that requires **evidence in exchange for funding**. Funding contracts often explicitly require (or **carry an expectation of**) a program or project evaluation. (Fletcher and Dyson, 2013, p.421, **emphasis added**)

The issue of expectation in evaluation has increased in light of recent research; particularly in emphasising the role of funders, as leaders. One leader always has expectations regarding their employees and hope these are met, however, the increasing popularity of evaluation as a funders’ compulsory practice has raised additional preoccupations on the match of their expectations.

Summing up, the role of the leader in power relationships in evaluation has been manifested in two dimensions: NPOs and funder agencies. NPO leadership is a social relation that tends to operate in a funder-recipient based context where an

imbalance of power is visible. The leadership arising from the funders' perspectives is concerned with the NPO's perception of funders' demand for evaluation, and the management of their expectations. This sub-section identified the need to further investigate how NPOs' leaders manage their organisations and how that management influences the organisation's EP.

2.5.3.3 Consequences of power relations in evaluation

Resistance

Resistance to engage in evaluation has been widely referred to in the literature on EP (Donaldson *et al.*, 2002; Geva-May and Thorngate, 2003; Taut and Brauns, 2003), yet its discussion has been focused, primarily in relation to the utilisation of evaluation results and less so, on the practice of evaluation (an exception is Saunders *et al.* (2011), as detailed in chapter four).

Resistance has been defined as “*human behaviour aiming to maintain the status quo, in the face of some form of real or perceived pressure (through evaluation) to change this status quo*” (Taut and Brauns, 2003, p.248). This definition shows that resistance arises from power relations, manifested through a mandatory evaluation, and in that sense, is a consequence of the attempts made to maintain the (power) status quo. The greater the distance between practitioners and their managers' concept of the social reality, then the greater will be practitioners' resistance to evaluation (Carter, 1971). In the same vein, Scriven (1991) notes the connection between resistance to evaluation and the generated fear, particularly, fear of losing funding, job, and prestige; and in general, fear of losing self-esteem and power.

The analysis of resistance, as a major problem for the practice of evaluation, has yet to be investigated in-depth (Datta, 2001). Existing studies have attempted to explain the origins of resistance, so that possible explanations pave the way towards a deeper understanding of the phenomenon within EP (Preskill and Torres, 1999; Abma, 2000). For example, Preskill and Torres (1999) note that because the goal of evaluation is to contribute to the learning of practitioners and organisations, a visible change has to occur – learning outcomes have to be met and evidence has to be demonstrated – consequently, people resist. Besides, the goal of evaluation in contributing to learning contains two problematic assumptions: first, it assumes that practitioners consciously want to learn from evaluation¹⁰, second, it assumes that by learning through evaluation a visible change occurs. Questions could be raised about how visible is this change or how practitioners occupying different positions across the organisation depict that change (Saunders *et al.*, 2005, 2011). The extent of their resistance could be illustrative of their perception of evaluation.

Taut and Brauns' (2003) values-pluralist's position explain resistance by acknowledging the human dimension in evaluation. Their study addresses possible psychological explanations for resistance that occurs throughout the evaluation process and addresses how evaluation influences the individuals in their social context, by exploring three explanatory factors for resistance to evaluation:

¹⁰For a review of professional learning communities, see Stoll *et al.* (2006).

1. Evaluation itself (divergence around the approach, method and interpretation of evaluation results);
2. Evaluation context (particular features of the broader context);
3. Personal features (practitioners' values and expectations, as well as, personal characteristics of the evaluator).

Taut and Brauns' (2003) study argues that evaluation, its context and personal features are all relevant to explain resistance. The key interest of their work to this thesis is their focus on the particular role of practitioner, regarding the resistance to the evaluation, particularly when most studies focus on the evaluators' role.

Judgement

One of the problems of the performance-based approach to evaluation is its pluralistic standard for judgement, which means that when a social intervention is evaluated, a multitude of opposing stakeholders' views are present (Greene, 1999). Moreover, what one considers 'quality' or 'good' in relation to a given intervention changes, as our mind evolves, so performance approaches to evaluation fail to capture these critical dimensions of program quality, which are central to EP. The direct association of evaluation with judgement is also noted by Taut and Brauns (2003):

A negative emotional state, often triggered by evaluation because of its association with "doubting", testing and **judgement, has a negative effect on co-operative behavior.** (Taut and Brauns, 2003, p.254, **emphasis added**)

Taut and Brauns' (ibid.) association of evaluation with judgement finds the ideal terrain in contexts where power relations operate. In this respect, the issue of

judgement has also been explored in the literature, from a realist approach, through the idea of evaluation as assisted sense-making (Julnes and Mark, 1998). This idea considers the role of evaluation as a contribution by “*extending, enhancing, and checking*” (p. 35) the natural sense-making that practitioners are involved in, regarding social programs and practices. Mark and colleagues (2000) place evaluation practice in the broader context of human judgement as an aid; so that evaluation is implicitly recognised as an informal human judgement, with or without a formal aid. As Smith and Brandon (2008) argue:

Teachers, students, parents and other **stakeholders do not hold off on making evaluative judgements about educational programs and practices until the “evaluator” arrives.** (Smith and Brandon, 2008, p.59, **emphasis added**)

The main challenge faced by this realist-based approach to evaluation is that the informal judgements within EP tend to occur across a variety of circumstances and contexts. Some may have support from a professional external evaluator, whilst in others cases, the judgements are intrinsically linked to a context of power relationships. As a result, the conditions in which these judgements occur strongly shape the way practitioners attribute worth to a project, as it happens within NPOs in the social domain (Carman, 2007). Furthermore, the defensive posture taken by practitioners to evaluation seems to relate to their fear of being personally judged (Saunders *et al.*, 2011), which may connect to the difficulty in separating the project’s evaluative judgement from their own personal judgement.

Scholars from neuroscience studies advocate that thinking about evaluation implies making judgements and that is harder than making descriptions, because

it involves more parts of the brain (Langer, 1997). This difficulty may lead to emotional distress, low self-esteem and what Donaldson *et al.* (2002) term as “*evaluation anxiety*”(p.261). Their work examines some of the negative consequences that can lead to evaluators’ anxiety, for example “*the lack of access to important information and data; compliance and cooperation problems; false reporting; effects on bias and validity; and reduced utilisation of evaluation findings*” (ibid. p.261). Despite the relevance of the strategies explored to prevent evaluation anxiety, Donaldson *et al.*’s (2002) work exclusively focuses on the role of evaluators, thus it would have been interesting to know their position about whether practitioners also developed this anxiety alongside their differences and similarities.

Ownership

Previous studies suggest that some NPO practitioners have not been consulted on the process of evaluation decision-making; subsequently, they seem to resist participating in evaluation (Taut and Brauns, 2003; Saunders *et al.*, 2011). Consequently, practitioners’ sense of not owning their evaluation tends to increase (Greene, 1988), particularly within a funder-recipient based context, where power relationships dictate working priorities. Hence, power relationships are not only relevant to EP but to the empowered practitioners (Gong and Wright, 2007).

The issue of ownership in evaluation connects with aspects of participation and empowerment, implying a values-pluralist theoretical standpoint. In the tradition of empowerment evaluation, ownership appears as a natural consequence of giving the necessary skills to put people “*in charge of their own destinies*”

(Fetterman, 1994, p.4); empowering their competence in EP. The contrary propels the idea of the evaluator as a detached expert, which makes practitioners “*dependent on an outside agent*” (ibid.) lacking competence, self-determination, and self-esteem. Taut and Brauns (2003) also refer to the lack of trust when practitioners are left aside during the evaluation decision-making process:

If in addition, evaluation is being determined **without the consultation of programme practitioners**, they will infer a lack of co-operation on the part of those contracting the evaluation. Such a **pervasive breakdown of trust** can result in all involved parties being in a vicious circle. (Taut and Brauns, 2003, p.254, **emphasis added**)

The ‘pervasive breakdown of trust’ was also mentioned by Schwarz and Struhkamp (2007) in two case studies, on evaluation in Germany’s HE reform, which explored the connection between evaluation and trust. Their intention was to gain understanding about how mechanisms of trust (and mistrust) are linked to evaluation, and in doing so, different meanings are assigned to evaluation:

What was considered ‘evaluation’ depended on the rules and contingencies of (group) dynamics and **had evolved and changed over time**. (Schwarz and Struhkamp, 2007, p.330, **emphasis added**)

Insights from the case studies suggest that trust probably “*got lost along the way*” (ibid. p.333) during the EP and this finding prompted some of my research observations to explore how evaluation is practised in a DE domain. Schwarz and Struhkamp (2007) argue that evaluation can either *build* or *destroy* trust; but in their study, evaluation does not seem to help building trust, so ambiguity arose. Despite the thought-provoking title, “*Does evaluation build or destroy trust?*”, their work could have pinpointed other effects of trust in evaluation, such as modification or improvement, alongside the dichotomy build/destroy, as

it would have been interesting to follow those additions in other domains. Their analysis of rich data seems conditioned from the beginning by two propositions – either evaluation builds or destroys trust – which may have limited further fresh insights on the relationship between evaluation and trust.

From the studies mentioned thus far, it surfaces that the notion of ownership in evaluation connects with notions of trust and mistrust, as a consequence (and symbol) of the power relationship and its arrangements within a funder-recipient based context.

2.5.4 The role of practitioner in evaluation practice

Many scholars have debated about the practitioners' challenge to practice evaluation, in relation to the role of professional evaluators, as experts (Abma, 2000; Saunders, 2006; Hart *et al.*, 2009), so that the relevance of the practitioners' role in managing power relations emerges as requiring more attention. The term 'practitioner' has been used in the literature on EP to refer to program managers and practitioners in general (Myers-Walls, 2000; Saunders, 2000), despite the reduced number of studies illuminating only aspects of practitioners' role (see Appendix 3; glossary). Although, there has been little substantive research into the ways in which practitioners give meaning to evaluation, Taut and Brauns (2003) consistently explore resistance and judgement by placing practitioners at the heart of EP, and Seppanen-Jarvela (2004) found that managers tended to be negative and critical towards evaluation, though they acknowledged its significance. One aspect that stands out from the literature is the practitioners' attitudes towards evaluation, and that seems critical to the thesis' focus.

2.5.4.1 Practitioners' attitudes towards evaluation

A key concern in the literature of EP is to understand the centrality of evaluation in practitioners' everyday practice (Carman, 2007; Saunders, 2000; Saunders *et al.*, 2011), as illuminated by Saunders (2000):

In most cases, **evaluation is not the central core of their [practitioners] working practice**. They are usually called upon or think it useful to undertake evaluations from time to time. In most cases they are managers, administrators, nurses, and teachers etc., i.e. individuals working with a wide variety of professional preoccupations. As such, they will have the 'practicality ethic' of people working in organizational settings. (Saunders, 2000, p.14, **emphasis added**)

However, a gap in the literature appears between the assumption that practitioners are knowledgeable evaluators (Carman, 2007) and the statement above that, in fact, they are ordinary people, for whom evaluation is not a core practice (Saunders, 2000). To elucidate further, more in-depth case studies looking at practitioners' attitudes towards evaluation, are needed.

While the key role of practitioners in their participation in EP, generally recognised especially when contrasting with evaluators' practice (Ayers, 1987; Feuerstein, 1988; Greene, 1988; Abma, 2000; Hart *et al.*, 2009), some scholars suggest a "*co-discovery*" technique, as a compromise to reduce the gap between the expert evaluator and the practitioner (Mayer, 1996, p.75). Other commentators have demonstrated an additional challenge for practitioners in responding to a multi-funder type of requirement (Hoole and Patterson, 2008), which is when a single intervention is funded by various funders; thus data has to be gathered, analysed and used for more than one evaluation. To overcome this challenge, practitioners tend to "*slice and dice data any number of ways to*

meet funders' requirements"(2008, p.98). Hoole and Patterson's argument that evaluation, as presently practiced, fails the needs of the social domain joins Greene's (1999) previous critique of the performance-based approach to evaluation; this critique is relevant to the current thesis, because it relates to practitioners' past experiences of (a possible performance based) evaluation (Owen and Rogers, 1999).

It has also been said that the meaning that practitioners give to evaluation should be taken more carefully into account in the evaluation, because their understandings affect how they practise evaluation (Seppanen-Jarvela, 2004). For instance, in exploring how NPOs evaluate their interventions, my research considers how practitioners' previous experiences of evaluation have shaped their current practice, and for instance, examines the "*sceptic factor*" in their attitudes towards evaluation (Drewello, 2001 cited in Taut and Brauns, 2003).

2.6 Summary and conclusion

The purpose of this chapter was to understand the current and future debates about EP as a research domain. The chapter contextualised the domain of EP, through the review of its theoretical landscape, to conclude that evaluation practice has been, and still is, an emergent research domain. It reviewed a number of theoretical calls that have strengthened the position of EP in the literature, and sheds light on a gradual tendency to the 'humanisation' of evaluation, through looking at what individuals do on a daily basis (practice), rather than, considering the performance of their interventions. As this thesis' research context is non-profit based, the chapter also focused on the preoccupation faced by such organisations as the DE in this case study, to

evaluate their social interventions. Equally, four main research debates in the field of EP were examined to generate the following relevant insights:

1. The implications of the influence of context in the EP are yet to be investigated in-depth.
2. There is a present need for alternative approaches for NPOs to evaluate social interventions, because a performance-based approach to evaluation has been demonstrated to be inadequate to attribute worth to these interventions; this was visible in a specific domain, the DE non-profit sector.
3. Evaluation, when practised in a funder-recipient based context where imbalanced power relationships occur, influences how it is perceived.
4. EP is presently understood as a mandatory request from funders, which practitioners have to fulfil, regardless of their knowledge base. However, it has been implicitly assumed that practitioners and their line managers have sufficient knowledge to meet these funders' request.
5. The role of practitioners within a funder-recipient based context appeared under-researched in the literature on EP.

The gap found in this chapter was that in-depth analysis of NPOs practices, their influences and how their EP changes over time have not yet been investigated as the main research focus, nor explored from a SPT view of evaluation. This chapter's argument is for the need of an in-depth exploration of EP and its change over time. To pursue this exploration, the literature states that a SPT framework is the most adequate approach (Saunders *et al.*, 2011). The following chapter advances the thesis in the direction of finding the reasons for exploring

evaluation from a SP approach, and also reviews the possible avenues, within that approach, to explore how evaluation is practised in a DE context. As the review suggests, this exploration should look at what individuals do on a daily basis – their practice.

This chapter suggested that a performance-based approach to evaluation has failed to attribute the worth of social programs, so that a shift is needed in placing practice at the core of evaluation processes (Saunders *et al.*, 2011). The main argument of the chapter proposes that the shift is yet to come, especially in a DE context. In progressing with the review of the literature, chapter three reviews the theories of practice, focusing on how a social practice theory has evolved to support the understanding of the contemporary social reality.

3 THEORIES OF PRACTICE: A SOCIAL PRACTICE PERSPECTIVE

3.1 Introduction and structure of the chapter

This chapter outlines social practice theory (SPT), as conceptualised by Reckwitz (2002) and Shove *et al.* (2012), and situates it within theories of practice.

The chapter starts by outlining the origins, contextualisation and elements of a social practice, through a review of the past contributors to this domain. It then moves on to the analysis of the trajectories of practice in order to depict change, the formation of practices and how they connect, co-occur and change over time. It does so by mapping differences, similarities and limitations in these theorists' frameworks. The chapter ends with the main conclusions from the review of the literature on SPT. The overall structure of the chapter is represented in Figure 3.1 (overleaf).

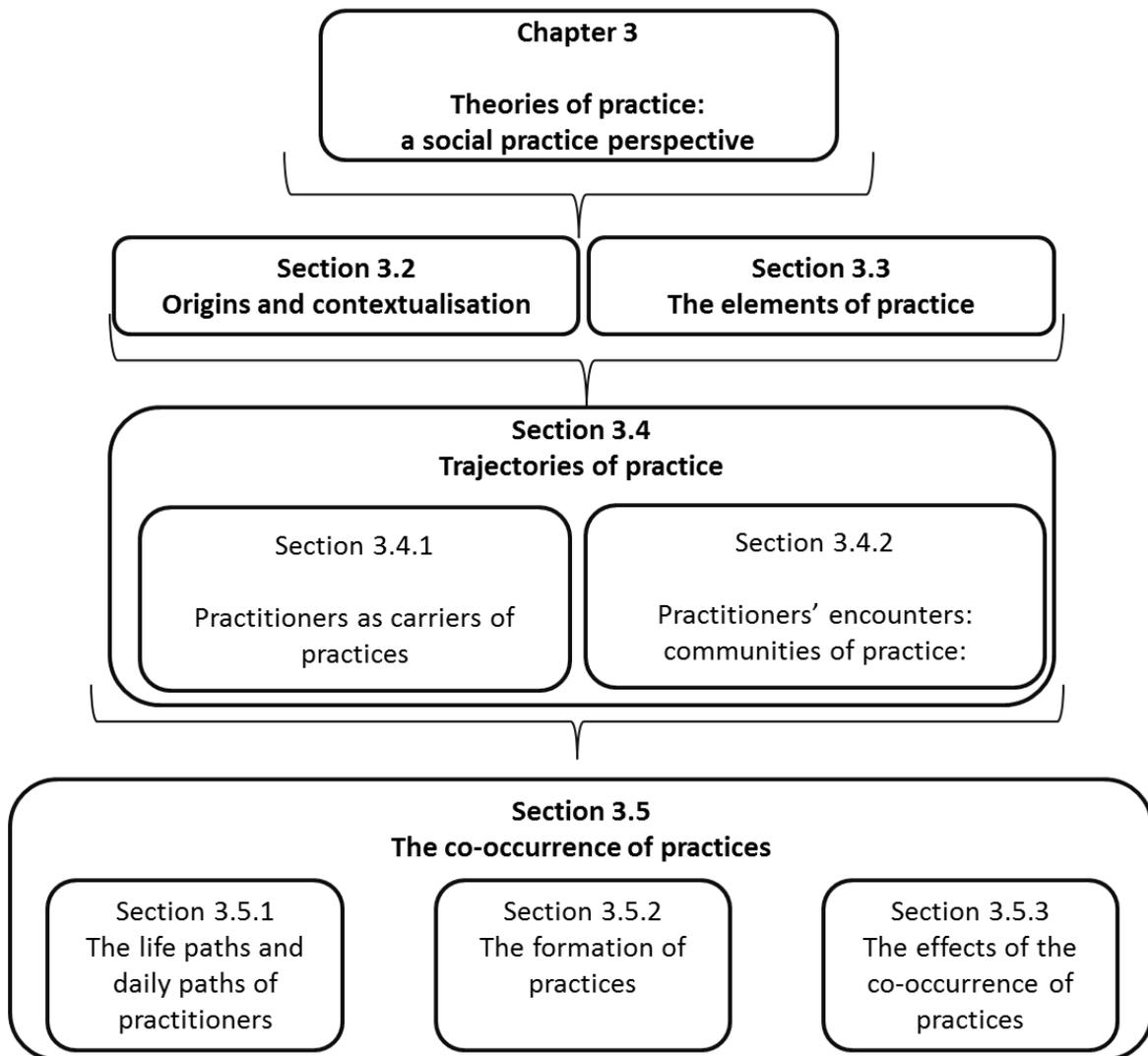


Figure 3.1: Structure of Chapter Three

3.2 Origins and contextualisation

This thesis adopts SPT to explore how evaluation is practised in a DE context. The main reason to adopt SPT instead of other branches of social theory is the aim to explore what people do when they evaluate social interventions - their practices: not their mental qualities (Schutz, 1972; Lévi-Strauss, 2004 [1962]); not their discourse (Foucault, 1972; Geertz, 1973); or even their interactions (Habermas and Burger, 2008). Mentalism, textualism and intersubjectivism are the other branches of social theory that are out of the scope of this review, because these appear to focus the investigation of social reality in other aspects of practice (mind, discourse and interactions).

The interest in social theory has its origins in the late 1960s and early 1970s and has been developed and influenced by other theorists, for example, with Foucault's (1972, 1990) analysis of relations, between the elements of bodies, agency, knowledge and understanding; Bourdieu's (1977) work on outlining a theory of practice and Giddens's (1984) theory of structuration, as a version of practice theory. In the domain of empirical anthropology and sociology, the contributions of Garfinkel's (1984) ethnomethodology; Butler's (1990) gender studies and Latour's (2010) science studies have gained momentum within the family of social theories. Since then, contemporary research has advanced and applied a framework of the dynamics of social practice (DSP) in domains, such as science, technology and society (Pred, 1981, 2007; Shove, 2007; Shove *et al.*, 2009), change behaviour (Shove, 2003; Shove *et al.*, 2012) and climate change (Shove, 2014). The commonality across the original studies on social theory, until today, is the pursuit for an in-depth exploration to understand social reality.

Whilst some of these theorists privilege some specific elements, such as discourse, agency and the intersection of structures to understand social reality (Foucault, 1972; Garfinkel, 1984; Habermas and Burger, 2008), others have chosen to look to peoples' behaviours to depict their understanding of the world (Giddens, 1979; Schatzki *et al.*, 2001; Reckwitz, 2002; Shove *et al.*, 2012). A SPT view was advanced by Reckwitz (2002) as a conceptual alternative that places the unit of analysis in 'practices', as a "*form of bodily activities, mental activities, 'things' and their use, a background knowledge in the form of understanding, know-how, states of emotion and motivational knowledge*" (ibid. p. 250). Accordingly, a practice represents a "*pattern which can be filled out by a multitude of single and often unique actions reproducing that practice*" (ibid.). Reckwitz (2002) illustrates what a practice is by giving the example of a certain way of consuming things that can be filled out by several actions of consumption. Another example more relevant to this research is that a certain way of attributing value and worth of social interventions can be filled out with a myriad of evaluation approaches. Social practice theory (SPT) explores what people do, what are their routines and behaviours and "*how a certain nexus or interconnection of these actions, behaviours and routines – practices – affects the way social activity is undertaken in a specific place and time*" (ibid. p. 258).

More recently, Shove *et al.* (2009; 2012) examined the dynamics of practices through depicting their interactions in the same setting and the influences shaping the persistence of one practice, at the expense of another practice's change. Although they draw on SP theorists reviewed here such as Reckwitz (2002), they apply a series of "*simplifying moves*" (Shove *et al.*, 2012, p.23) to summarise the elements of practice. Previous studies that employed a DSP lens

in other domains (Shove, 2003, 2007; Shove *et al.*, 2009) have shown the importance of depicting what constitutes a practice, because it allows for the identification of the elements of practice needed, which are outlined next.

3.3 The elements of practice

An initial conceptualisation of SPT by Reckwitz (2002) identifies seven elements of practice: body, mind, things, knowledge, language, structure and agent. Shove *et al.*'s (2012) conceptualisation of practice in their DSP framework draws on Reckwitz (2002), in terms of ontology and epistemology of the social sciences; however, it differs regarding the categorisation of these elements. Firstly, they interpret practice through three elements: material, competence and meaning to understand why people do what they do on a daily basis - see Table 3.1(overleaf). Secondly, to describe how social practices persist or change, a DSP view examines the co-occurrence of a given practice with other ongoing practices.

Description	Elements of practice by Reckwitz (2002)	↔	Elements of practice by Shove <i>et al.</i> (2012)	Description
	Bodily performances and emotional activities	Body	}	
Social relation between mind, body and objects	Things			
Mental activities (desiring something; understanding the world; knowing how to do)	Mind	}	Meaning	Social and symbolic significance of participation at any one moment
Ways of wanting, feeling, avoiding unwanted things; motivational knowledge	Knowledge			
Exists only in its routinized use	Language/discourse	}	Competence	Multiple forms of understanding and practical knowledgeability; knowing in the sense of having the skills required to perform (deliberate cultivated skill)
Structured set of routines (temporality, repetition and social reproduction)	Structure/process			
Carriers of practice, depending on past experiences; use of know-how as motivational knowledge	Agent			

Table 3.1: Frameworks of elements of practice by Reckwitz (2002) and Shove *et al.* (2012)

As Reckwitz’s (2002) categorisation of elements of practice has informed Shove *et al.*’s (2012) recent framework, I take the latter to review the elements of practice. This DSP approach accepts that there are some elements that need to interact in order to designate a working activity or other task, as a practice or “*enactment*” of that practice (Shove *et al.* 2012, p.23) in which the elements of material, competence and meaning have to intersect. In an evaluation context, material is interpreted as things, objects, and tools through which meaning is assigned to a certain performance (practice); competence is interpreted as having the skills required to participate and conduct an evaluation; and meaning is interpreted through visions, ideas, or aspirations of the importance of

participation at any moment across a project's evaluation. Next, I briefly describe these elements of practice and their relevance for this research.

- Material

The intention of the use of a DSP framework is to explore how links are made between the elements from which a practice is composed. The role of 'things' is instrumental in the understanding of how people relate to objects, thus, existing research has recognised them as inherently linked to and intertwined with objects (Schatzki *et al.*, 2001). In the same way, Gherardi (2009) advances this idea by adding that objects and their materiality can be depicted as "*materialised knowledge*" that probes people and relates with them. There is now an established consensus around the idea that things, as well as objects, are an element of practice (Røpke, 2009). In my research, the materiality of EP is explored through progress reports, meetings plans, and events flyers, alongside the observation of the physical space dedicated for the practice of evaluation.

- Competence

Shove *et al.* (2012) combine notions of "*know-how, background knowledge and understanding*" to resume those as "*competence*" (p.23), despite their concern about different perspectives on knowing, within SPT. They aggregate these previous notions with skills and techniques to emphasise the ability "*to participate with the requisite competence in the complex web of relationships among people, material artefacts and activities*" (Gherardi, 2009, p.118). Shove *et al.*'s (2012) notion of competence is the sense of having the skills required to participate, implying knowing and understanding; for example, through the skills needed to conduct an evaluation. Although Gherardi's (2009) point is relevant,

Shove *et al.*'s (2012) framework of DSP is favoured to explore EP and change over time because it primarily focuses on its elements and how they change across a three-year project's timeline.

- Meaning

Mental activities, emotion, and motivational knowledge (Reckwitz, 2002) can depict the meaning *through the social and symbolic significance of participation at any one moment*" (Shove *et al.*, 2012, p.23, **emphasis added**). Shove *et al.* (2012) condense Schatzki's (1996) theorisation of past and future of meaning, emotion, and motivation by including "*symbolic meanings, ideas and aspirations*" grouped as meaning (Shove *et al.*, 2012, p.14). In analysing and discussing these terms, Schatzki (1996) focuses on "*teleoaffective structures*" (p.88), which means embracing the "*ends, projects, tasks, purposes, beliefs, emotions and moods*", as critical to organising and locating social practice within a specific setting – "*timespace*" (ibid. p. 89). Shove *et al.* (2012) discuss this relatively new concept to elaborate on the idea that what people do has a previous account ("*history*") and a situated background ("*setting*") (ibid. p. 24). They draw upon Schatzki's (1996) term of "*timespace*" (ibid.) to suggest that what people do is future-oriented, therefore both aspects ("*history and setting*") tend to be visible during the occasion of the '*practice*'. However, the relevance of the meaning - as an element of EP - for this research, is to explore whether Shove *et al.*'s (2012) perspective of an evaluation 'timespace' is visible in a DE domain. Summing up, a DSP framework requires that practices are represented by interdependent associations between their elements – material, competence and meaning – that are "*out there waiting to be linked together*" (ibid. p.24).

3.4 Trajectories of practice

Early conceptualisations of SPT have addressed the issue of change in relation to how practice theory influences the social reality (Reckwitz, 2002). Specifically, previous theorists found out that “*people are unknowingly engaged in reproducing and enacting multiple and varied cycles of change, simultaneously shaping the lives of practices and being shaped by them*” (Shove *et al.*, 2012, p.77). As a result, the use of practice is employed as the smaller unit of analysis in examining that social reality. However, the obvious consequence is that what people do on a routinised basis is sometimes unintentional, so their perception of what they do can often be different from what they actually do. A SPT lens depicts this difference by following a person, event or project’s evaluation trajectory over a period of time; and maps their continuum of practices. This aspect is fundamental to my research, because it allows for the contrast of participants’ articulations of their perceived EP with in-depth observations of what they actually do in their everyday routine. To depict these differences on the trajectory of participants’ EP it is relevant to use diverse data collection methods. Shove *et al.* (2012) refer to “*trajectories*” (p.19) to conceptualise how the elements of practice - material, competence and meaning - interact. They suggest that by following a trajectory of a routinised practice, it would be possible to depict change in that practice, as well as within its elements. This distinction is of paramount importance to this research, as it allows the deconstruction of EP to examine each element; and in doing so, novel insights may explain the contemporary practice of evaluation in a DE setting, as well as shed light on how (and why) it changes over time. As a result, a key aspect of these trajectories of practice is the relationship between their elements, as explained:

If specific configurations are to remain effective, connections between defining elements have to be renewed time and again. This suggests that stability and routinisation are not end points of a linear process. Rather, they should be understood **as ongoing accomplishments in which similar elements are repeatedly linked together in similar ways.** (Shove *et al.*, 2012, p.24, **emphasis added**)

It would be possible to affirm that either a “*practice*” (links between elements being made), a “*proto-practice*” (links between elements not yet made), or an “*ex-practice*” (links between elements no longer being made) (Shove *et al.*, 2012, p.25) is taking place within a specific context of EP.

3.4.1 Practitioners as carriers of practices

Whilst for Reckwitz (2002), a practitioner is “*a carrier of practice who carries certain routinised ways of understanding, knowing-how and desiring, as well as many different practices which need to be coordinated with one another*” (p. 250), others have recently referred to practitioners as individuals who carry practices, regardless of their professional affiliation, competence, or power (Shove *et al.*, 2012). The way practitioners, as carriers, understand, know, desire, avoid, or want, are features of their singular participation in a given practice and not qualities of their individuality. As a result, the issue of participation mirrors how practitioners carry a specific practice over time. The impact of practices on peoples’ lives is closely related to the “*social and symbolic significance of participation*” or in DSP terms to their “*meaning*” (ibid. p.65). As in many other practices, the issue of participation relates to the numerous and different opportunities people have to participate or not; therefore, the importance of past experiences is critical for the inequities of access and participation, as well as

central for what might happen next. Individuals are “*constantly taking up and dropping out of different practices, as their lives unfold*” (ibid.)

Subsequently, an important assumption of the DSP approach is that the delineations of any practice – “*where it is reproduced, how consistently and for how long*” (ibid. p. 63) – depend on the role of practitioners as carriers. In other words, practitioners are the human agents that bring a practice towards an enactment zone; they do it in their contextual settings, with their own values, personal beliefs, and tensions with funders (e.g. evaluation practice), which in turn, shape how these practices are carried and, ultimately, persist or disappear. Two important issues for my research arise from this assumption: first, how EP attracts DE practitioners and how they spread through social communities; second, how DE practitioners encounter practices and then become their carriers. What has to be in place for practitioners to become carriers of practices? Is it a conscious movement? After following some carriers’ trajectories, Shove *et al.* (2012) proposed that as the carriers’ commitment develops or decreases “*some practices become more deeply anchored and embedded in society while others disappear*” (ibid. p.64). One of the strengths of a DSP approach to my research is its historical recollection of a multitude of other practices, and working activities, to analyse how EP changes, persists, fades away, or otherwise over time; how it is carried, and how it interacts with other concurrent practices or activities in a given context (further detailed in section 3.5; co-occurrence of practices). This leads to a second assumption about social relations, which is that the practices are stable and it is argued that to analyse change, one has to assume that the practice “*is stable, or at least, stable enough for one to tell whether someone is involved in them or not*” (Shove *et al.*, 2012, p.66). Some ideas from

Lave and Wenger's (1991) challenge this point of stability, particularly that practitioners encounter the practice, i.e., when they have to carry evaluation forward to the next stage of the process or to another stakeholder. Also, the sequence involved in carrying a practice varies from one practice to another, so what applies to the practice of learning may not apply to the practice of evaluation. The next section examines how practitioners encounter practices and the context of that encounter.

3.4.2 Practitioners' encounters: communities of practice

Some literature on communities of practice (CoP) and learning strongly connects with that of SPT, for example, Shove *et al.* (2012) outline various ways practitioners start to carry practices, bringing with them the dimension of learning to make sense of how social relations emerge. For the purpose of this research, I review the dimension of CoP and discard further studies on social learning (Barton and Tusting, 2005; Akiba and Alkins, 2010; Wenger, 2015), because that would open another realm of literature for the focus on EP.

In attempting to understand how practices are conceived, shared and carried across domains, a DSP lens unpacks the concept of CoP to conclude that community and practice constitute each other (Wenger, 1998). While Wenger and Snyder (2000) define CoP as "*groups of people informally bound together by shared expertise and passion for a joint enterprise*" (p.139), Shove *et al.* (2012) maintain that:

the ties and connections through which **practices** develop and circulate, and by means, of which **they reach and capture new recruits [practitioners], do not necessarily map onto**

organisational or institutional structures. (Shove *et al.*, 20102, p.67, **emphasis added**)

It is not only the practitioners' present experience and their organisational structure that counts, but the links which are formed through, and as a result of past and present experiences. Turner (2001) emphasises the notion of "*shared practices*"(p.120), focusing on how they come to be shared and regarded; and two issues that emerge are whether the community is characterised by the practices in which their members engage and what happens when individuals engage in multiple practices (see section 3.5.3; the effects of the co-occurrence of practices). Shove *et al.* (2012) propose that multiple communities arise as links "*overlap and extend beyond the margins of any one practice*" (ibid. p.68). As a result, accidental encounters and unpredictable episodes tend to occur, as practices are constantly (consciously or unconsciously) being shared. It is in this context that the notion of trajectory of practice is paramount to represent the dynamic link between the practitioners and the EP. In this respect, a trajectory of EP enables the examination of how the elements of material, competence and meaning intersect and change over time; through capturing what has happened in, and over a specific period of time. However, the analysis of this dynamic link between practitioner and practice has challenges for their social life, particularly that of becoming a "*full practitioner*" (Lave and Wenger, 1991, p.60), which means that a practice can be carried out at any moment by practitioners with distinct levels of commitment and expertise; with some drop-outs along the trajectory. The practitioner's encounter may have a "*transformative effect*" as persons and practices "*change, re-produce and transform each other*" (ibid. p. 68).

The consequence of practitioners becoming committed to the practices they carry out is that they may change and become what they do (Becker, 1977); for example, one practitioner that has responsibility for producing evaluation reports may want to have a say in the way data is monitored and then organised in that report. Similarly, a practitioner may tend to become what they think about (Nightingale, 2013). For example, someone who has confidence in their ability to produce a thorough evaluation report becomes someone who can do just that; or someone else that frequently thinks that funders may judge their evaluation becomes someone who lacks confidence. This suggests that the meaning of evaluation can be altered through practitioners' thoughts, choices and commitments and in becoming what they do. In DSP terms, how practitioners “*choose and commit*” (Shove *et al.*, 2012, p. 69) to the practices is critical to understanding how these practices persist or disappear, and it is equally fundamental to analyse how practitioners carry some practices and disregard others (*i.e.* how they continually participate in one practice and probably, consciously, resist another). What is already known from previous research, is that by the time practitioners interrupt this travelling in which they are engaged, that practice is no longer sustained (Schatzki *et al.*, 2001; Reckwitz, 2002; Shove *et al.*, 2012) - as the link between their elements is broken. The process of carrying practices forward and, or, across domains is meaningful for its trajectories, as a pattern emerges:

At any one moment, a practice consists of a **composite patchwork of variously skilled, variously committed performances enacted and reproduced by beginners and by old-hands alike**. Patterns of career development are, in combination, relevant for the trajectory of the practice as a whole. This is so that newcomers and those with more experience

inevitably reproduce somewhat different variants. (Shove *et al.*, 2012, p.71, **emphasis added**)

For example, when novices become experienced-workers, their skills (competences) circulate not among themselves, but according to how their relations are structured in the organisation; shaping how trajectories of practices are developed. Hence, it is helpful to think about practitioners' social responsibility in the "*persistence or disappearance*" of a practice (ibid. p.73), and how their competence to the (re)configuration of the social world is influential.

In summary, before a DSP lens, participation in a given practice is mainly perceived as practitioners' engagement or lack of it, but does not go beyond that factual event, either they participate or do not. From a DSP lens, to participate in a practice enhances the meaning assigned to that practice, leading to a new sphere of articulations about social responsibility in carrying and sharing the practice along the way. The contrary is also true - the lack of participation in that practice assigns the symbolic meaning through which a practice is understood, communicated and carried (or not) by practitioners. In this line, the notions of the practitioner as a 'carrier of practice', and the 'trajectories of practice' are central to exploring how EP changes over time. For the purpose of this exploration it is also relevant to understand how practices are formed, co-occur and the aspects that influence, change, or alter their co-occurrence over time.

3.5 The co-occurrence of practices

As the focus of my research illustrates, the practice of evaluation occurs alongside other working activities that DE practitioners recurrently do; for

example, management, decision-making, negotiation, and even cleaning. This section illustrates how this co-occurrence of evaluation practice with other working activities is likely to influence, and change, evaluation practice conducted by DE practitioners.

3.5.1 The life paths and daily paths of practitioners

Central to the DSP framework is the idea that the “*life paths*” (Shove *et al.*, 2012, p.77) of practitioners interconnect with, and combine to constitute the life paths of individual practices. For the purpose of the thesis, I borrow this term because it contains the meaning that the trajectory of EP has to be looked after by practitioners and other actors, socially responsible for the wellbeing of that practice. Other possible terms, such as ‘life course’ or ‘life course of practice’ appear more generalist, but may obfuscate the close relationship with an active practitioner. Personally, the latter term implies that a practice follows its course (in a forward motion), regardless of any intentional action.

In brief, the life path of a practitioner closely influences the life path of their individual practice. In taking this articulation to practitioners, within this research, interesting findings across the trajectories of their EP may arise, particularly the notion that taking one path rather than another configures opportunities for the future. Therefore, in an evaluative situation in which practitioners move into institutional roles, their “*daily path*” (Shove *et al.*, 2012, p.77) ends up structured by priorities and aspirations that are relevant for the articulation of their whole life path. This may explain the reason practitioners hold so tight to past professional experiences, because their daily path has been structured by previous institutions (in some cases with opposing visions) and

their practice has been shaped within that context, resulting in a fixed imprint that persistently resists change. An illustration from the literature on evaluation is when practitioners learn useful mechanisms that generate a successful evaluation, they are likely to adopt that “*embodied experience*” (ibid. p.69) and make it dominant, (sometimes unconsciously). For instance, previous studies illustrate that the EP appears dominant when a final evaluation report is completed (Saunders *et al.*, 2011) and a revealing insight is whether that dominance is shared by the responsible practitioner or by others.

The contribution of SPT to the exploration of (evaluation) practice does not lead directly to the alteration of power structures, but can allow novel insights into understanding how those structures are created, maintained or altered (Shove *et al.*, 2012). Although Reckwitz (2002) might have implicitly hoped that by looking at practices, some inequalities would be altered, this was not an explicit intention. In contrast, Shove *et al.* (2012) illustrate that a DSP lens allows for understanding policy and problems, because it permits practitioners to realise “*how policy agendas and problems are defined and framed*” (Shove *et al.*, 2012, p.140). In turning to EP, practitioners have the opportunity to understand how and why certain funders’ requests are structured by a formal and performance-based approach and how practitioners can engage with funders’ rationale regarding the reasons for requesting evaluation in a specific (prescriptive) way rather than another. Practitioners can also frame evaluation from this practice-based view ensuring a beneficial trade-off for them, while responding to funders’ evaluation requests. However, whilst practitioners are re-framing their EP or unconsciously being dominated by a funder’s power structure, the life path

of their organisation carries on with a multitude of other practices (co-) occurring.

3.5.2 The formation of practices

Theories of practice have particular relevance in the formation of practice, principally in the way the elements of practice interact, balance, or disintegrate (see section 3.3; the elements of practice). Whilst some theorists focus on how other branches of SPT operate, in relation to the formation of practices (Reckwitz, 2002), Shove *et al.* (2012) provide an in-depth examination of the implications of the co-occurrence of practices, for their elements and for the wider context. SPT assumes that a practice involves the active integration of the elements - material, meaning, and competence - and through these is seen “*as a provisionally recognizable entity*” (ibid. p.82). This assumption leads to another, which assumes that practices are defined by what practitioners actually recognise them to be, at a given moment. The life path of a practitioner strongly shapes the life path of their (evaluation) practice, but because of the dynamics of everyday life, that practice is provisional. This aspect of a provisional practice resonates with that of “*provisional stability*” developed by Saunders *et al.* (2005, p.37), in relation to how an evaluation message travels across a specific trajectory of implementation. This notion appears to have enhanced the practice of evaluation within complex, popular, and politicised domains of HE and the European Union (EU) funding programs (further developed in chapter four; a SPT view of evaluation). However, the way EP is framed is relatively consistent and the permanent integration of its elements is vital to the formation and interaction of evaluation with other practices occurring simultaneously –

concurrent practices. In other situations, a single practice needs to be deconstructed and its elements purposively separated to be taught to novice practitioners, as the example of RUFDATA (Saunders, 2000) illustrates (see section 4.3; SPT applications to EP).

This research is interested in observing if and how EP is communicated, inducted and exemplified to practitioners, by whom, and with what intention. To do so, I explore the co-occurrence of other practices or working activities restricting EP; as well as new practices replacing those already in existence. This replacement, however, may be gradual or non-existent, because some older practices still exist but are enacted less frequently than usual. For example, when practitioners formulate interventions to tackle a specific problem by means of designing a project, the practice of design is vital, whereas when the project's evaluation begins, the practice of evaluation tends to replace the practice of design (EuropeAID, 2004). In this vein, Lizardo and Strand (2010) suggest that the knowledge within the formation of each practice is "*born of first-hand, embodied experience and, does not live in the realm of discursive consciousness*" (p.11), which means that there are challenges for the EP, because a situation of concurrent practices does not allow for the embodied experience to persist (in terms of knowledge and competence); rather it facilitates the accumulation of practices, instigating potential change. In this regard, it is acceptable, and almost inevitable, that a practice is likely to concur with others, generating consequences for their enactment and for the practitioners involved. To better understand the importance of the co-occurrence of practices and if and how it influences the change over time, I outline three important aspects: first, the pre-formation, formation and de-formation of

practices; second, the role of time, and, third, the role of space in the formation of practices.

3.5.2.1 Pre-formation, formation and de-formation of practices

Shove *et al.* (2012) conceptualise the pre-formation, formation and de-formation of connections between practices, to analyse how they are connected and shape each other. Three scenarios were drawn: the first, in which practices exist without being integrated, because their elements are not in competition; e.g. evaluating and cleaning, as in the activity of cleaning the office, practitioners use different skills from those used in their EP; second, practices are provisionally linked by co-occurrence or co-dependence; e.g. the practice of project design and project evaluation, because one implementation cannot be evaluated if a project has not been designed (EuropeAID, 2004). Third, the connections between practices are no longer sustained; e.g. the practice of reflection and the practice of evaluation in a DE setting are challenged by the funder's evaluation demand, which restricts the possibility of practitioners' reflective practice (Bracken and Bryan, 2010; Bourn, 2014). As we live in an interdependent world it might appear that connections between practices are obvious but they "*can and do co-exist and co-depend in ways that their practitioner-carrier barely recognises*" (Shove *et al.*, 2012, p.84). Considering a given practice through this analytical lens helps the process of recognition and, perhaps, increases understanding of why people do things in a certain way, rather than another.

3.5.2.2 Time and practice: timeline as a tool to depict change

In describing the trajectories of change, Shove *et al.* (2012) outline how a certain practice has changed and become established in society and how its embedding

has influenced the structure and social arrangements of that practice. The potential of the analysis of the trajectories of practices is that by analysing the past, one can make sense of a specific trajectory of practice; why its meaning has changed and what type of competences are currently required that were not necessary in the past. As a result, interpretations of how a certain practice's position has emerged, sustained or disappeared over time are crucial to explore its trajectory of change; particularly, how the notion of time interferes with changes in practice (over time). For instance, while some commentators conceptualise time through the relations between time given and payment (Hochschild, 2000), others highlight the multi-tasking idea and conceptualise the day as a series of short and repeated episodes (Hess, 2006). A DSP lens draws on Southerton (2003), who suggests that these same short episodes are the outcome of changing interactions between practices. Practices come with sets of requirements necessary for competent and meaningful engagement (Hess, 2006), which have consequences for how the time is organised and scheduled.

A timeline, as a tool to depict change in EP, enables the reflection about how the social world is fast changing; hence, one implication of the co-occurrence of practices for change is the need to capture which practices or working activities are persisting and why. Also, a timeline (of EP) makes visible other persistent reconfigurations of practice, as well as depicting the effects of concurrent practices, in terms of time, space and resources. Hence, the relevant issue about time, to my research, is the demands that a specific EP make in terms of their duration, timing and sequence, so topics, such as how the research participants respond to the requirements of EP set by the funders; how the current evaluation practised by NPOs is becoming more demanding; how the change in meaning

associated with the time contours of EP is likely to be investigated. These topics are explored through the analysis of the co-occurrence of EP with other working activities, in a DE working environment.

3.5.2.3 Space and practice: working environment

There are various ways in which space represents and influences relations between practices. Shove *et al.* (2012) considers that places like offices and homes can have “*emergent consequences for the trajectories of individual practices and, hence for the collection of practices that are, and that are not, enacted in such environments*” (p.85). As mentioned throughout this section, it is likely that a routinised practice clashes with other practices or working activities, however this is not negative or positive, it is the way DSP operates, because spatial and temporal aspects do impact on how various practices co-occur. A DSP lens offers a conceptualisation for this problematic in terms of “*complexes and bundles [which are] loose-knit patterns like those based on co-location, sometimes turn into stickier forms of co-dependence*” (ibid. p. 87). Nevertheless, this is not always the case; there are many situations in which practices co-occur well. “*Complexes*” are practices that end up depending on each other, either because of “*timing, proximity or necessary co-existence*” (ibid.). These complexes are likely to appear in various working environments and my research explores how the co-occurrence of EP with other working activities influences the practice of evaluation, in a DE context. This leads to how practices co-occur through four possible effects: competition, collaboration, selection and integration.

3.5.3 The effects of the co-occurrence of practices

3.5.3.1 Competition and Collaboration

The way that practices connect with each other assumes their conditional existence upon the manifestation of the three elements – *material, competence and meaning* – and that practices compete to find a carrier practitioner. Their argument assumes that a certain practice needs the existence of other practices to be enacted. On the one hand, certain practices generate elements on which other practices depend; so this sequence implies dependence and further collaboration. For instance, in the current research the practice of evaluation depends upon the design, formulation and submission of a given project for funding, so that an evaluation is subsequently requested. On the other hand, whilst practices do compete for elements, they are united by those elements that they share. Again, in the same example, EP and design share the same human resources, competence and space settings. As Shove *et al.*, (2012) note, practices become connected with each other, as each practice shares certain elements to sustain each other, whereas when practices compete for certain elements, they break the links that connect them and tend to disappear. It does not always seem easy to depict whether the interactions between practices are competitive or collaborative. As the focus of my research illustrates, the practice of evaluation occurs alongside other working activities that DE practitioners recurrently do; for example, to manage the strategic future of a NPO, or to decide on their next funding bids. In this example, the practice of evaluation and the activity of management compete for the time allocated to a single practitioner, so that a decision has to be made about whether this practitioner practises the evaluation or collaborates with managing the organisation strategic future. The practice of

evaluation and this working activity may compete for their elements; and possibly the latter may reinforce the connections needed to improve the practice of evaluation. If this improvement flourishes, collaboration between practices and other working activities may occur. In collaborating or competing for resources and attention, some practices can obtain dominant status through selection and integration, as explained next.

3.5.3.2 Selection and Integration

Shove *et al.*'s (2012) ideas on the dominant practices are informed by previous STS studies, specifically, a model of innovation that suggests the existence of “*micro, meso and macro*” levels of interaction between practices when new arrangements emerge (ibid. p. 92). These categories of interaction interweave, so that the micro level informs the meso, which in turn “*structures and [...is] structured by macro-level landscapes*” (ibid.). For example, in an evaluation setting a micro level of interaction between evaluation and other working activities is represented by practitioners’ collaborative work within the YP, with a meso level being the NDEC’s practice of evaluation, in general, and a macro level, the funders’ requirements for evaluation. From one level to the other, the elements to be carried are selected, so the connections become gradually more intense and established. This means that change, in an evaluative context, may face resistance, possibly through adaption, overlapping or modification. However, the connections between EP and other working activities are not made through a linear route of co-dependence, but they rather include “*nodes, knots, relays and points of convergence and amplification*” (ibid. p. 94), which make the emergence of EP contingent on how other practices or activities interweave and not only on how they compete.

In the current research, an apparent node or knot may be represented by external actors coming into the evaluation, generating adaptation, or modification of how EP co-occurs with other working activities. The main aspect taken from this DSP literature to my research is that different influences shaping EP may unfold simultaneously, generating various effects of change across the evaluation timeline, such as competition or collaboration, through dominance; and selection or integration, through adaption, overlapping or modification.

3.6 Summary and conclusion

This chapter advocated a SPT view of evaluation that values multiple layers of reality, accepts complexity as a characteristic of a fast changing interconnected world, and celebrates practitioners' trajectories of (evaluation) practice and their change over time.

The purpose of this chapter was to provide an outline of the key concepts within SPT, as conceptualised, mainly by Reckwitz (2002) and Shove *et al.* (2003, 2007; 2012). While initially, the literature focused on Reckwitz's (2002) examination of the elements of practice, particularly, on how a SPT view of these elements differs from other SP branches' views; it also indicated Shove *et al.*'s (2012) preference to focus on the DSP, which emphasises how trajectories of practice can depict change; how practices are formed; and how they connect, co-occur and change over time. The review presented in this chapter diverges from Reckwitz (2002) and others' views, on purpose due to different research avenues taken. As an important addition to the understanding of SPT in contemporary research, this chapter highlighted that the issue of participation mirrors how practitioners carry a specific practice over time. In turning to my

research, the argument emerging from this chapter focused on the adoption of DSP (Shove *et al.*, 2012), as a suitable lens to investigate the phenomenon of evaluation and how it has been practised and carried by practitioners over time.

Collectively, the studies reviewed in this chapter outlined a critical role for 'practice' in researching what people do on a daily basis, which allowed me to conclude the importance of a SPT framework to chart the shifting nature and direction of the practice of evaluation. The next turn, and final chapter, in the review of the literature describes how scholars have applied SPT in their studies on EP.

4 EXPLORING A SOCIAL PRACTICE PERSPECTIVE OF EVALUATION

4.1 Introduction and structure of the chapter

This chapter explores the intersection of social practice theory (SPT) with EP, and presents SPT's distinct insight into an in-depth exploration of peoples' routines of evaluation, by capturing their intentional and unintentional tasks over time, some of which had not been perceived as evaluation.

The chapter begins by examining how previous studies in EP have mainly used Reckwitz's (2002) SPT framework, and how Shove *et al.*'s (2012) framework of dynamics of social practice (DSP) supports the in-depth exploration of the phenomenon of EP. The chapter also outlines previous studies in EP to identify some potential gaps in the literature. It ends with the formulation of the research questions and the main conclusions drawn from this critical analysis. The overall structure of the chapter is represented in Figure 4.1 (overleaf).

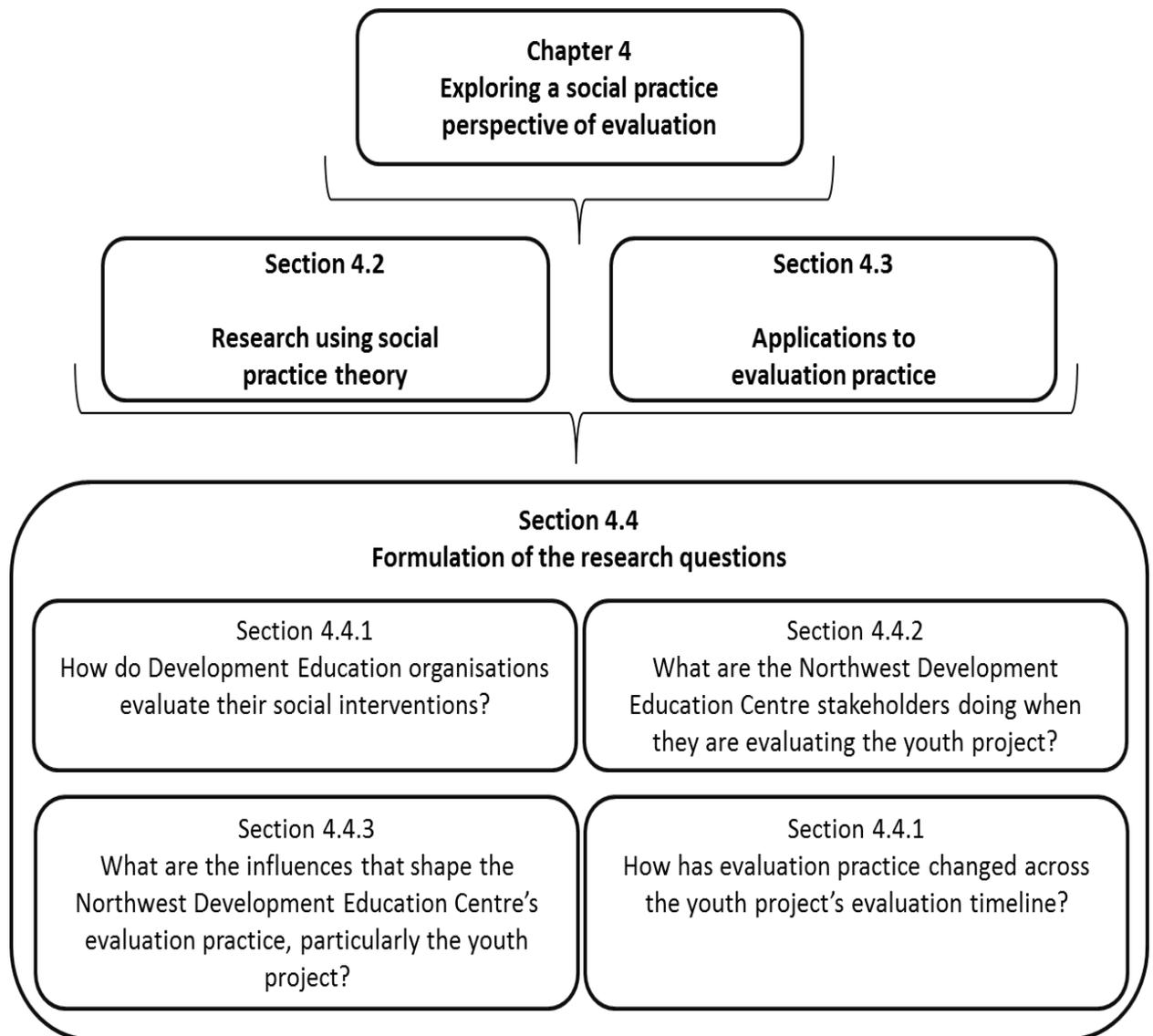


Figure 4.1: Structure of Chapter Four

4.2 Research using social practice theory

Chapter two notes the unquestionable importance of evaluation as a discipline, to improve the quality of the organisational interventions delivered, through learning the strengths and weaknesses of each particular case (Simons, 2009). Consequently, the evaluation process and its practice is particularly important in the context of the attribution of worth and merit of NPOs' interventions, particularly within a development education (DE) context, because these organisations are more likely to encounter difficulties gathering evidence of their activities, as they strive for social change. Chapter three concludes that a SPT lens challenges the traditional boundaries of evaluation, to discover novel ways of practice, particularly in relation to change over time and its effects.

This chapter reviews previous work grounded in practice-based evaluation, as an approach that values what people do, think, and what they assign meaning to, when establishing the value and worth of a social intervention (Saunders *et al.*, 2011). Some commentators reinforce this challenge, by advocating that the process of thinking about evaluation is, already, evaluation (Westley *et al.*, 2006). What are the consequences if thinking about evaluation is considered EP *per se*?

Practice-based evaluation is informed by a SPT view that privileges the contribution added to tackle a social problem, allowing for a variety of possibilities to attribute value and establish the worth of that social intervention (Fitzpatrick *et al.*, 2004; Saunders *et al.*, 2005; Saunders, 2012; McCluskey, 2011). A practice-based approach to evaluation may or may not be apt for the challenge of depicting unexpected changes in a globalised evaluation era, which

scholarly work emphasises is best done within a democratic vision of evaluation (Greene, 1999; Simons *et al.*, 2003). As such, a SPT view of EP is likely to contribute to the democratisation of evaluation, because it gradually alters the practitioners' perception of evaluation into a meaningful practice proposing other ways of measuring impact, rather than only carry out 'formal evaluations' (Parry-Crooke, 2014).

Previous works have reported the importance of a practice-based evaluation (Bemelmans-Videc *et al.*, 1998; Kushner, 2000; Schwandt, 2002, 2005), with empirical research focused on the potentialities of exploring practices by applying them to domains such as, European Union (EU) e-learning projects (EQUEL) (Bonamy *et al.*, 2004); Higher Education (HE) (Saunders, 2000; Saunders *et al.*, 2011; Trowler *et al.*, 2012; Trowler, 2013); EU structural funds programs (Saunders, 2011, 2012), and technology enhanced learning (Trowler *et al.*, 2014). The next sub-section reviews some of these empirical studies emphasising their important commonality of a SPT stance of evaluation.

4.3 Applications to evaluation practice

This section reviews Saunders' work of RUFDATA and further contributions to EU program evaluations (Bonamy *et al.*, 2004; Saunders, 2011, 2012). The reason for reviewing these studies is because of their common view informed by a SPT view of evaluation, inspired by Giddens (1976) and, later, by Reckwitz (2002). These studies were also applied in different contexts, so a critical review is important to define key messages and limitations in the current research of EP in a DE domain.

RUFDATA, which is a hallmark of Saunders' (2000) work, has been employed in other domains of social sciences¹¹ and stands for Reasons and Purposes; Uses; Focus; Data and evidence; Audience; Timing; and Agency. The context in which RUFDATA emerged was one where novice evaluators were being initiated into the process of evaluation, specifically, into the practice of planning the evaluation (Saunders, 2000). It is described as an approach to planning evaluation as a sequence of "*knowledge-based practices*" (ibid. p.13); referring to knowledge as one element of practice, visible through "*ways of doing things*" (ibid.). Additionally, it draws on the tradition of learning, from Lave and Wenger (1991), which resulted in Saunders' (2000) conceptualisation of evaluation as taking place within a community of practice. This context of depicting evaluation allows for a period of deep thinking, using enabling tools, such as reflexive questioning, particularly important to novice evaluators, with little or no previous experience, but likely to be charged with an evaluative task. This form of practising evaluation advocates a flexible approach, "*low in fidelity to give a sense of manoeuvre to practitioners*" (ibid. p.15), as well as the sufficient guidance for them to get evaluation "*off the ground*" (ibid.). RUFDATA is a situated approach that takes on board the influential role of context (Pawson and Tilley, 1997), the "*particularity*" of each intervention (Simons, 2015, p.173) and the uniqueness in practitioners' "*evaluative moments*" (McCluskey, 2011, p.103), which would serve the funders on a provisionally stable basis (Saunders, 2005).

¹¹ See Bamber *et al.* (2009) and Dickson (2011).

This approach is enacted through the formulation of some reflexive questions that form the basis of the EP. From a foundational perspective, the contribution of RUFDATA to research EP is unquestionable, however, Saunders (2000) seems to privilege the stage of planning and conducting an evaluation, rather than, looking at practice across the evaluation process as a whole. Although the planning stage is critical, the application of RUFDATA or some of its principles, across the whole process of evaluation could possibly enhance its EP, as well as benefit its wider application in other research domains. In terms of my research, this leads to questions about how the funders' request for the evaluation, the leadership of the organisation, and the existence of other practices and working activities, influence EP across the trajectory of the whole project.

Another important work is a case study of a European funded initiative, *EQUEL* (e-Quality in e-Learning), that analyses how a practice-based evaluation has been used in a complex multi-partnership context of an e-learning project (Bonamy *et al.*, 2004). The case study indicates the diversity of positions each group or individual occupies within the project and the subsequent complexity arising from the intersection of their positions, discourses and perceptions. *EQUEL*'s framework combined Saunders' view (2000) on practitioners' understanding of the program circumstances, with Fullan's (2011) acceptance of diversity and complexity as an *a priori* proposition in every process of change. While Bonamy *et al.*'s (2004) contribution suggests that it is necessary to support practitioners in sharing their questions and observations; it recognises that this support may, or may not be during the lifetime of the project. That seems to be a very critical aspect for my current research in EP within NPOs, as

most of their evaluations tend to operate within unpredictable circumstances of an external funding context.

Despite its use of a SPT view of evaluation framework, Bonamy *et al.*'s (2004) analysis would have benefited from further analysis about the factors influencing change, the effects of the co-occurrence of other practices, issues of sequence, time and space in the trajectory of practice, alongside their implications for the wider context. This could also be because it is prior to Shove *et al.*'s (2007; 2009; 2012) conceptualisation of DSP.

Moreover, Bonamy *et al.*'s (2004) study emphasises the aspect of learning in evaluative processes and how the notions of “*provisional stability*” and “*evaluation as bridging tool*” (Saunders *et al.*, 2005, p.39) may enhance EP. They recognise the nonlinearity of EP, through the critique that a metaphor of ‘implementation stair case’ appears to suggest a certain stabilisation of EP within a project’s implementation; see also Reynolds and Saunders (1987).

It would have been interesting though to see identified some connections to allow the exploration of these notions in other domains of social sciences. Also, EQUEL’s study (Bonamy *et al.*, 2004) re-frames Hall and Loucks (1978) seven stages of continuum from awareness to refocusing, regarding teachers’ stages of concern in project involvement (when dealing with new ideas), by simplifying it in three – exploratory, consolidation, and development - and this move has potential to be further applied and conceptualised in other domains, as it captures a broader sense of participants’ experiences of project implementation in evaluation.

In another research study about evaluation of interventions, policies and programs, Saunders (2011) outlines a practice-based approach to evaluation to capture the effects of EU interventions, particularly regarding the diversity and the complexity of the EU context. A fundamental feature of a practice-based approach to evaluation is the need for a shift from an attribution to a contribution perspective of outcomes achievement (Mayne, 1999). An attribution perspective contemplates the causality of outcomes achievement, due to a specific implementation or activity, whereas a contribution perspective accepts a provisional and yet stable input or practice as a relevant contribution towards the achievement of the desirable outcomes. In this evaluation of an EU program's effectiveness, Saunders (2011) calls for a greater attention on the role of practices as a preferable vehicle to capture the effects of interventions; especially within a funder-recipient based context. The practice approach is attached to *"whatever it is that people do as a response to a policy, an intervention, opportunity or initiative, and this constitutes an inductive method"* (ibid. p.93). By using a SPT lens, informed by Reckwitz (2002), Saunders (2011) sheds light on the risk of an excessively logical view of understanding change, dominated by a predetermined way of thinking. Nonetheless, in a situation of understanding change, neither practitioners on the ground, nor the policy makers, *"act in a rational way"* with established goals in mind (ibid.). People's practice of evaluation is sometimes unintentional and so a contribution approach to evaluation celebrates this possibility, whilst noting that some intentional influence occurs. Whilst Saunders' study was situated in a complex area of EU project funding, it would have been interesting to explore how a SPT contribution to evaluation applies in other domains of social sciences.

A final study considers the use and usability of evaluation as practices in two dimensions; first, how evaluation is used by real people in real time, and second, how the design of evaluation influences its use in a wider sense. Saunders' (2012) argument is that the impact of an evaluation can be traced back in relation to the changes in its practice over time. Furthermore, he argues that it is carried by practitioners over time, and that a SPT view adds value to evaluation in three dimensions: first by depicting use and usability as practice, practitioners strengthen their knowledge resources to make sense of evaluation. Secondly, by adopting a SPT view he suggests that use of evaluation is translated through the change of its practice, to the extent to which, and in which domains it is used or not used. Besides, Saunders (2012) challenges the view about "*what might be the nature of the knowledge resources available from an evaluation that creates the momentum for change (i.e. for it to be used)*"(ibid.p.426). Thirdly, a SPT approach enables the identification of practices of engagement to depict patterns of change over time. One thread of Saunders' (2012) argument lies in the fact that "*it is difficult to get practices to change*" (ibid.). My research draws on this argument to explore how evaluation is currently practised in a DE domain.

However, the assumption that practices are difficult to change is challenged from the DSP view that change is inevitably reproduced and enacted by people, regardless of their awareness (Shove *et al.*, 2007; 2009; 2012). A further thought leads to a dual-dynamic, one of a conscious and intentional change being difficult, with another unconscious and unintentional change happening in peoples' trajectories of life. Still, Saunders' (2012) analysis may have benefited from considering the added value of DSP to analyse use and usability of

evaluation, for example, whether evaluation has been used in the same ways or how change has been manifested in the way evaluation is used.

Shove *et al.* (2009; 2012) emphasise the use of the elements of practice to depict the dynamics of change over time and in simplifying Reckwitz's (2002) theorisation of seven elements to three – material, competence and meaning – (see Figure 3.1; the frameworks of the elements of practice), they were able to explore one element in-depth at a time, as well as how the others interwove. This process enables the capture of change in practice over time.

In sum, empirical and theoretical evidence strongly suggests the further application of a SPT approach in evaluation studies. Previous studies have demonstrated the usefulness of SPT in a number of dimensions. For example, in emphasising the core benefits of reflexive questioning in the planning of evaluation (Saunders, 2000); the recognition of a nonlinearity of EP (Bonamy *et al.*, 2004); the benefits of a contribution approach to better understand change in evaluation (Saunders, 2011), and the relevance of individuals' engagement to describe patterns of change in evaluation (Saunders, 2012).

However, these important studies have drawn upon a SP view (Reckwitz, 2002) that has been recently extended and updated by Shove *et al.* (2007; 2009, 2012). As illustrated in chapter three, Shove *et al.*'s (2012) particularly emphasises a number of dimensions that are likely to impact in EP such as the co-occurrence of practices and how the effects of this co-occurrence influence and change overall practices. With notable exceptions (Saunders *et al.*, 2011; Trowler *et al.*,

2012), these important issues have only been addressed superficially in evaluation theory and therefore remain an important gap to develop.

Moreover, while studies drawing upon a SP view have offered in-depth insights into how evaluation occurs in a number of contexts, repeated calls have been made for further studies that explore EP in other domains. For example, Saunders *et al.* (2011) encourage scholars to “*develop a meta-framework that yields the possibility of comparative research into evaluative practices into other social policy domains*” (p.226). In line with this call, the current study begins to address these gaps by extending a SPT view of evaluation into the domain of Development Education (DE).

The next section summarises the gaps presented in the three chapters of the literature review, as well as devises this study’s research questions.

4.4 Formulation of the research questions

The main argument emerging from the three chapters of literature is that a practice-based approach is likely to benefit the evaluation of social interventions, because it allows practitioners to reflect on the trajectory of their practice. This routinised reflection may lead to a transformation of practitioners’ perception of evaluation.

Chapter two debated that a performance-based approach to evaluation has failed to attribute the worth of social programs, so that a shift is needed to place practice at the heart of evaluation (Saunders *et al.*, 2011; Bourn, 2014). Chapter three outlined the theoretical foundations of SPT and examined the importance of capturing the change in EP over time.

Two frameworks were examined to conclude that the DSP approach (Shove *et al.*, 2012) allows capturing how a practice evolves, re-adjusts and modifies, through the analysis of the trajectories of practices, alongside the depiction of the practitioners' daily paths in becoming committed carriers of practice. Chapter four critically examined how scholars have applied SPT in their empirical studies on EP.

Table 4.2 (overleaf) illustrates the overview of the gaps identified, which informed the research questions devised. The following sub-headings explain the rationale of each research question.

Research questions	Type of gap	Gap in the literature (expanded)	Location in the thesis
Overarching RQ. How do Development Education (DE) organisations evaluate their social interventions?	Theoretical /Analytical	1. Inadequacy of the performance based approach to evaluate social interventions within a NPOs context.	Chapter two Sections 2.5.1 and 2.5.2
RQ 1. What are the Northwest Development Education Centre (NDEC) stakeholders doing when they are evaluating the youth project (YP)?	Theoretical	2. Current literature on SPT has not empirically analysed the dynamics of the trajectories of practices (formation, co-occurrence and its implications for the wider context).	Chapter two Sections 2.5.3 Chapter three Sections 3.4 and 3.5 Chapter four Section 4.3
RQ 2. What are the influences that shape NDEC's evaluation practice, particularly the YP?	Analytical	3. Prior empirical research conducted on how NPOs evaluate their interventions, has mostly quantitatively analysed large and medium size organisations, within popular and politicised domains of social sciences	Chapter two Section 2.5.2
RQ 3. How has evaluation practice changed across the YP evaluation timeline?	Theoretical/ Analytical	4. Updated and extended theoretical framework of SP offers potential new insights to important issues in EP (e.g. co-occurrence of practices) Inconsistent understanding of change in the EP, because a DSP lens has not yet been used to investigate how practices change in evaluation, more specifically within a domain of DE.	Chapter four Section 4.3

Table 4.1: Overview of the research questions and their rationale

4.4.1 How do Development Education organisations evaluate their social interventions?

The first theoretical gap concerns the present understanding on the theoretical approach to evaluation that best fits the social domain. While this review suggests several attempts that have conceptualised alternative ways to practise evaluation in a social domain (Fetterman, 1994; Greene, 1988; Saunders, 2000; Mertens, 2001; Patton, 2010; Bamberger and Segone, 2011), empirical evidence proposes an insufficient level of understanding on how NPOs evaluate their interventions (Greene, 1999; Carman, 2007). Furthermore, the literature notes that this gap is extremely prominent in the context of a DE research setting (Bourn, 2014). Additionally, literature on EP in NPOs appears contradictory; in some it is assumed that practitioners are prepared to evaluate social interventions, in accordance with funders' performance-based requirements (Carman, 2007; 2009), while others suggest that these organisations still struggle to find the adequate resources to undertake a performance-based evaluation (Greene, 1999; Bourn, 2014). This contradiction emerges within a DE domain, because most organisations deal with social and attitudinal change, and are asked by funders to evaluate their interventions from a performance-based approach.

Although recent commentators recognise the need to shift from a performance to a practice-based approach to evaluation (Bourn, 2011, 2014), there is scant literature on how DE NPOs evaluate their interventions. The analytical gap of knowing how these organisations practise evaluation generates the overarching

research question that my study investigates: how do DE organisations evaluate their social interventions?

4.4.2 What are the Northwest Development Education Centre stakeholders doing when they are evaluating the youth project?

The second analytical gap concerns an overall inconsistent understanding of EP in the social domain. The literature notes that the way EP is understood is highly shaped by many aspects, of which context (Chouinard, 2013) and power (Atjonen, 2015) stands out. Whilst current literature advocates that a powerful relationship from funders towards practitioners seems to jeopardise the practice of evaluation (Greene, 1998; Schwandt, 2005), most empirical studies on EP place either the evaluator or the funder at the heart of the research (Henry and Mark, 2003; Carman, 2007; King, 2007; Gong and Wright, 2007). As few studies have focused on the practitioners' perspectives of evaluation, with notable exceptions of (Greene, 1988; Taut and Brauns, 2003; Seppanen-Jarvela, 2004), there is a calling for further research exploring their role in EP, as well as their routines and practices. To examine their role this review indicates that there is also slim literature about what individuals are doing when evaluating DE projects, such as the YP. Therefore, a potential gap of looking at evaluation from a SPT view was identified in the literature, alongside the under-researched role of practitioner, as a key player in a funder-recipient based context. These gaps generated RQ1 that asks what the NDEC stakeholders are doing when they are evaluating the YP. The aim is to cover the SPT aspect of EP indirectly rather than introducing the term 'social practice' into the question. The main reason for this decision is that when disseminating results I want to share this thesis with

my research participants and, possibly, with a wider audience, so I want people to understand the nature of my study and what has shaped my motivation. Thus, I decided to keep the questions simple, rather than populate them with theoretical terms. However, I have ensured that the theoretical rationale is present and each question has its place in this research.

4.4.3 What are the influences that shape the Northwest Development Education Centre's evaluation practice, particularly the youth project?

A third analytical gap concerns the design and scope of the empirical research conducted on how DE NPOs evaluate their interventions. Past research has mainly adopted a quantitative or mixed-methods approach to analyse large and medium size NPOs, within popular and politicised domains of social sciences, such as healthcare (Carman, 2007), teaching (Buckley *et al.*, 2015) and governance (Berner and Bronson, 2005; Alaimo, 2008). A related potential gap identified the need for more empirical research to investigate how small organisations evaluate their social interventions, and as a result, this research qualitatively explores how evaluation is practised within the relatively recent and a-political domain of DE; asking what are the influences that shape NDEC's EP, particularly the YP?

4.4.4 How has evaluation practice changed across the youth project's evaluation timeline?

In recent advances in SPT, Shove *et al.* (2012) developed a framework of the DSP, which investigates specific features, such as how practices change over

time; the co-occurrence of practices and the effects of change in practice and their implications; these features had been disregarded by previous social theorists such as Reckwitz (2002), and not yet substantially applied in empirical studies on EP. My study covers this gap by conducting an in-depth exploration of EP within a DE context, from a DSP view. That gap generated RQ4, about how EP has changed across the YP evaluation timeline.

In sum, Table 4.2 resumes the theoretical and analytical gaps, alongside the respective research question that emerged from the literature review. Research questions 1, 2, and 3 revisit the overarching question about how DE organisations evaluate their social interventions. Having recognised that the current articulation of the research questions is mostly derived from the literature, I acknowledge the influence of past experiences, theoretical views and fieldwork in the final development of these questions in chapter five, the methodology chapter.

4.5 Conclusions

The purpose of this chapter was to critically examine how scholars have applied SPT in their empirical studies on EP noting the specificities of Reckwitz's (2002) and Shove *et al.*'s (2012) strands, as well as the limitations arisen from that application. The chapter acknowledged the main limitations of the use of previous versions of SPT in evaluation, arguing that there is room for exploring EP through the analysis of its dynamic of change over time. From a SPT view, the close investigation of peoples' practice means that a wider array of actions, activities, attitudes, informal conversations, silences, and absences will be captured, which forms the practice of evaluation (Saunders, 2011).

This chapter suggested the potential of a SPT perspective to research evaluation, because of the features of capturing change over time enable the researcher to trace back why certain things occur or not in NPOs' social interventions. The chapter also mapped some of the empirical SPT research applied to EP to conclude that there is a potential gap in the application of DSP, in the domain of EP; and specifically in a DE domain of EP.

Although the studies presented in this chapter are a key contribution to the research on EP, some potential research gaps are summarised in Table 4.2 and were used to generate four research questions. The apparent inadequacy of the performance based-approach to evaluation in a DE setting, combined with a few qualitative in-depth studies in EP, generates the overarching RQ 1: How do DE organisations evaluate their social interventions? Surprisingly, few qualitative in-depth studies have empirically analysed what people do as a routine when they evaluate. This gap also relates to the under-researched role of practitioner as

a key player in a funder-recipient based context. As the current development of DSP is yet to be applied to empirical studies on EP, this potential gap in the literature generated the RQ 1: What are the NDEC' stakeholders doing when they are evaluating the YP? The previous research conducted on EP was a quantitative and mixed-methods analysis of large and medium size NPOs, within popular and politicised domains of social sciences, which led to ask RQ 2, what are the influences that shape NDEC's EP, particularly the YP? In responding to a call to closely examine the role of practitioners in EP, this review noted an inconsistent understanding of EP in the social domain and a lack of research on its change over time, particularly in a DE domain; which generated RQ 3: How has EP changed across the YP evaluation timeline?

In this literature review I argued that the term 'evaluation' has rigid connotations when aiming to evaluate DE projects. In this respect, a SPT view of evaluation appears to accept each practitioner's set of skills as valuable and acknowledges the contingencies of a complex and fast changing world. My view also supports the view that what one mindfully does in a specific "*evaluative moment*" (McCluskey, 2011, p.103) is provisional and yet stable enough (Saunders *et al.*, 2011) to contribute to the merit of EP. The position coming from this chapter's argument, in combination with chapters two and three, is that each moment of EP counts; formal or informal; using arts-based or evidence-based methods can act as a "*bridging tool*"(Saunders *et al.*, 2005, p.38) in the evaluation process. Each moment of EP has its own legitimacy, its trajectories of change and its critical incidents over time, calling for deep reflection (Shove *et al.*, 2012). A practice-based evaluation celebrates the emergence of change; the embeddedness of that change in the contexts that it operates in order to

influence; a shared social agreement that permeates participation, responsiveness and values pluralism of the social reality, recognising the global challenge of the historical and cultural context of the setting. All of these live in a permanent transition, in an ongoing and never-ending dynamic of daily life.

5 METHODOLOGY

5.1 Introduction and structure of the chapter

The purpose of this chapter is to provide an overview of the research methodology used to investigate the practice of evaluation within an under-researched domain of development education (DE). The chapter starts by outlining my ontological and epistemological views in order to locate the research within a social constructionism paradigm. It also outlines the methodological choices regarding the research design, explaining the use of a qualitative case study approach and the appropriateness of an in-depth single case (section 5.2). Section 5.3 describes my approach to fieldwork, by discussing the practicalities of the pilot and the main study, as well as the data collection methods used. It also examines and describes the itinerary of data analysis stages, alongside outlining how the data is trustworthy and authentic in the context of this study. The latter part of the section details the ethical considerations and section 5.4 summarises the methodological and technical issues encountered, as well as the importance of a reflective space. Section 5.5 summarises the main conclusions and the methodological choices undertaken. The structure of the chapter is represented in Figure 5.1 (overleaf).

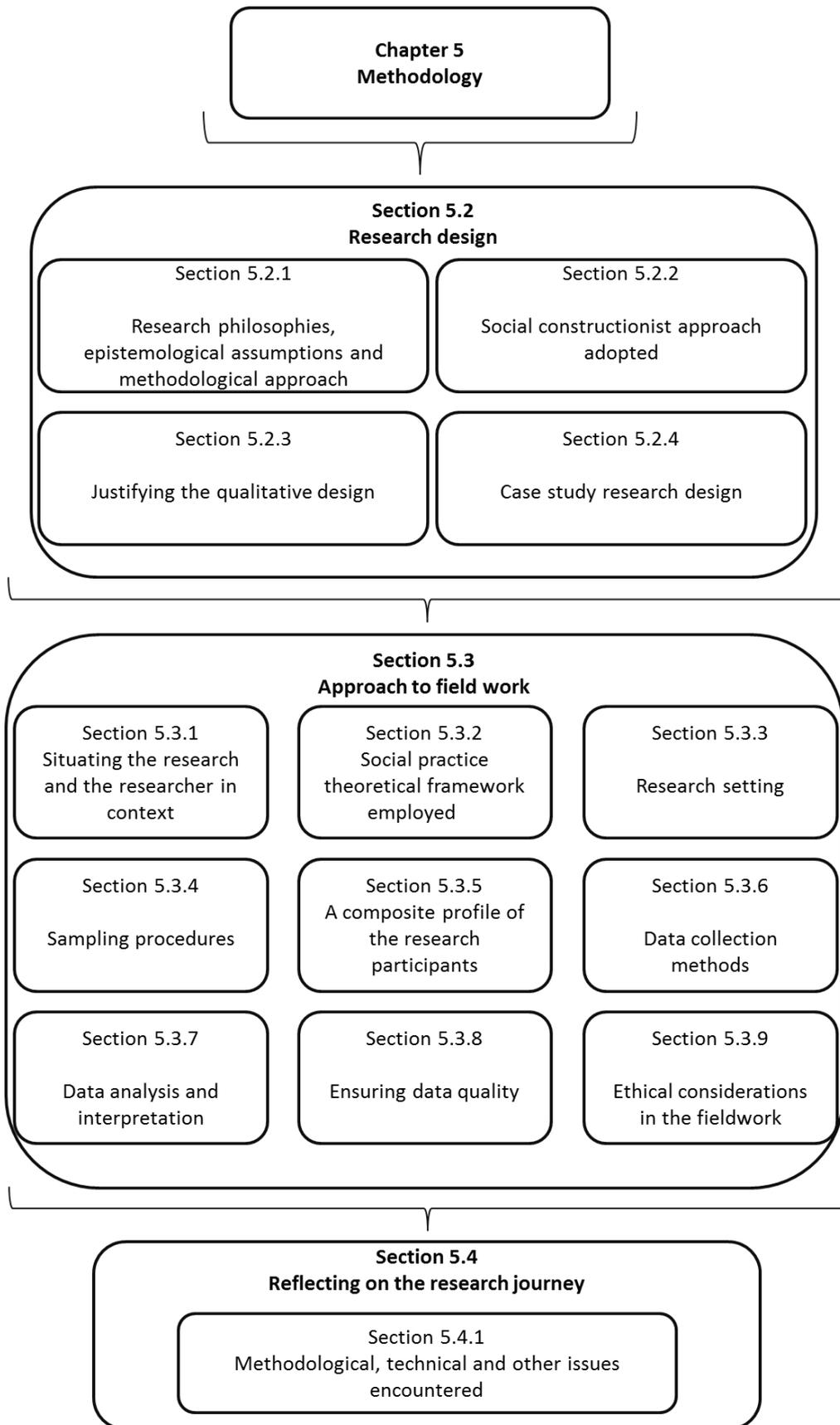


Figure 5.1: Structure of Chapter Five

5.2 Research design

The presentation of this thesis may imply that the process of finding a research approach, ongoing refinement of the research questions, revisiting of the methodology and methods and subsequent data analysis has been a linear process; however, there is no such a thing as linearity in qualitative research, rather there is messiness (Law, 2004). This has been a part-time study, a lengthy but enjoyable process, in which messiness has played a critical role. This section starts by presenting the research philosophy that support my theoretical approach throughout the thesis. Next, I justify the qualitative design used in the form of single case study.

5.2.1 Research philosophy, epistemological assumptions and methodological approach

This study investigates how evaluation is practised in a DE research setting, through an in-depth exploration of what participants do on a daily basis, regarding evaluation. As a researcher, I came to the conclusion that everyone has an ontological position, whether they realise it or not. As my research journey has been characterised by strong reflection, my ontological position is that knowledge is socially constructed through interactions with and through the observation of the actors involved. I take the philosophical view that there are no certainties in the social world, but social constructions as a set of assumptions informing the research. Therefore, in the current research, I am enquiring about the way evaluation as a social practice is constructed, and by whom.

When referring to epistemological assumptions, I share Crotty’s (1998) view that each epistemological position implies a critical difference in how we conduct our research, so I started with an everyday issue that needed to be addressed and subsequently planned my research journey. Although Crotty (1998) suggests that the four basic elements of any research process are epistemology, theoretical perspective, methodology and methods, my research process did not evolve in the same order, as illustrated in Figure 5.2 below:

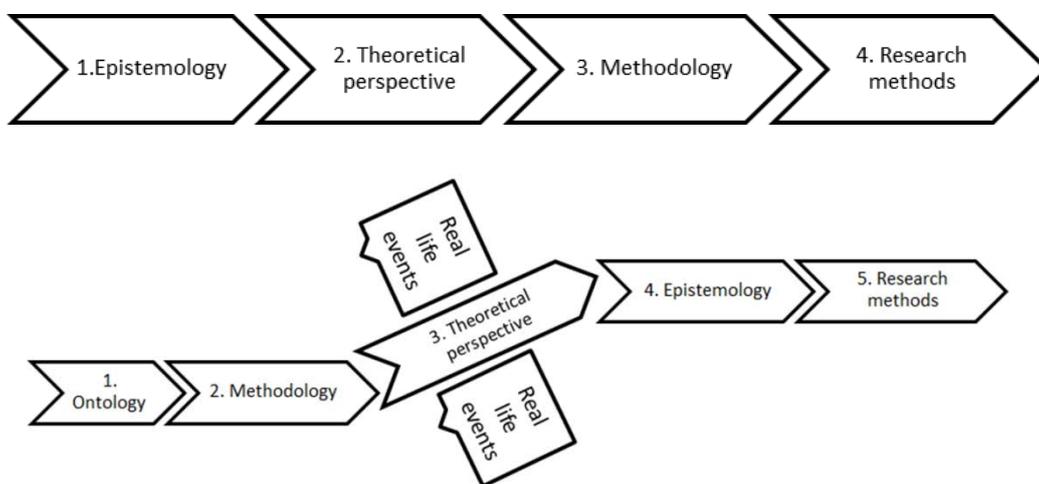


Figure 5.2: The order of research elements (adapted from Crotty, 1998)

For instance, from a social constructionist ontological view, the element of methodology appeared due to my previous involvement with the case organisation. During this period, as a volunteer, it became explicit that one single in-depth perspective would be relevant to enquire about participants’ experiences of EP. Therefore, as I started to read literature about evaluation as a social practice, the theoretical perspective emerged quite naturally. While some authors claim that research is not theory-free (Yin, 2008; Silverman, 2011), others advocate for a value-driven, opinion-driven, situated view that does not

contain any theoretical value or position (Creswell, 2007; Simons, 2009). When starting a research process, those value-driven positions are refined (and/or sometimes totally altered) into a theoretical perspective. It was this theoretical perspective on the real-life event observed, in combination with literature reviewed that supported the development of specific research questions. The research questions that incorporate the purposes of my research led to the appropriate methodological choices (Crotty, 1998), and a sense of epistemology appeared. Consequently, the natural choice of using longitudinal observational methods suggested that the sense-making process of the social world would be produced by experiencing one specific and time bound real-life event over a period of time.

Figure 5.2 above challenges Crotty's (1998) position on the linear order of the research elements, suggesting that there is room for others' journeys. I describe next the social constructionist approach adopted.

5.2.2 Social constructionist approach adopted

A paradigm or worldview is a basic set of beliefs that guide action (Guba and Lincoln, 1988). In examining the phenomenon of evaluation and its practice through the exploration of participants' experiences, this thesis adopts a social constructionist paradigm. In the research tradition, this paradigm may be labelled social constructivist, in which the meaning is constructed through the mind (Mertens, 2005), but for the purpose of this research, I will use 'social constructionism', because it takes the view that the meaning of the social world is constructed through practices (Reckwitz, 2002). Accordingly, to depict participants' experiences, it is critical to understand their values, meanings, and

beliefs, by examining their practices. A social constructionist researcher makes interpretations of what “*they see, hear and understand, [which cannot be detached from their own] background, history, context, and prior understandings*” (Creswell, 2007, p.21). As Gergen (1999) suggests, social constructionism assumes “*that for any state of affairs a potentially unlimited number of descriptions and explanations is possible*” (p.47); therefore, a social constructionist approach reflects my intention to explore in-depth how evaluation is practised in a DE setting. Also, this approach allows me to explore how reality is socially constructed by a multitude of people active in the research process (Schwandt, 2005).

5.2.3 Justifying the qualitative design

Qualitative designs have become established and accepted forms of enquiry in the domain of social science (Guba and Lincoln, 2005; Creswell, 2007; Silverman, 2011). Hence, the focus of this section is to justify why the qualitative approach was chosen, rather than to compare quantitative with qualitative approaches. Four different aspects of fit are proposed: ontological approach, nature of the questions, research approach, and contribution to the field.

First, the qualitative nature of this study fits well with my ontological approach on how knowledge is constructed. As Merriam (1998) points out, the key to understanding qualitative research lies in the idea that meaning is socially constructed by individuals in the interactions with their world. Second, some scholars have stressed that the choice of a qualitative approach should rely on the nature of the questions (Patton, 2002; Denzin and Lincoln, 2005; Creswell,

2007). Qualitative methods are particularly appropriate to provide “*answers to questions that stress how social experience is created and given meaning*” (Denzin and Lincoln, 2005, p.4). Creswell (2007) also proposes the use of qualitative research designs to gain in-depth understanding about a particularly complex phenomenon or process, which suits this study’s aim of exploring how evaluation is practised within a DE domain. Third, a qualitative case study research approach is a “*rigorous exploration of how the researchers’ values and actions shape data gathering and interpretation and how people and events in the field impact on the researcher*” (Simons, 2009, p.4). Due to immersion in the fieldwork, qualitative investigations produce rich descriptions of the processes and meaning for each of the participants in the research (Guba and Lincoln, 2005), and use in-depth description and exploration to privilege individuals’ lived experiences (Creswell, 2007). Doing qualitative research is in many ways no different than everyday life: “*it is complex and sometimes downright chaotic*” (Silverman, 2011, p.14). Lastly, the choice of qualitative approaches must make sense in the development of a contribution to the field, within a particular area of knowledge (Silverman, 2011). For this reason a qualitative in-depth exploration supports an investigation of how evaluation is practised over time in a DE domain; which DE literature considers under-researched (McCollum and Bourn, 2001; Coriddi, 2008; Liddy, 2010; Bourn, 2011). Thus, a qualitative study approach is an appropriate design to accomplish a deeper understanding of evaluation practice (henceforth ‘EP’) in a specific research setting of DE.

5.2.4 Case study research design

Case study is a study of the *singular, the particular, the unique* (Simons, 2009, p.3). The approach to this case study was mainly inspired by Simon's contribution to the evolution and practice of case study research in education and educational evaluation (Simons, 1987) and partially by Yin (2008), through the feature of an embedded case. This research adopts a case study approach, rather than a case study, only as a method. By case study approach, I mean "*an overarching research intent and methodological purpose, which affects what methods are chosen to gather data*" (Simons, 2009, p.3). The current research uses a whole case study approach because of its main purpose of exploring the particularity and the uniqueness of the EP in a specific DE setting – its distinctiveness. Moreover, this approach values multiple perspectives of participants, observation in natural daily-life circumstances, and interpretation in context. This approach is also consistent with my ontological and epistemological positions, in the way that participants socially construct the meaning and interpretation of the world, with my presence as a social constructionist researcher. Therefore, in this thesis the term 'case study approach' contains my ontological and epistemological considerations, the research design adopted, the methodology used, the data collected and the analytical framework. In addition, Simon's (2009) approach focuses on the use of case study within evaluation research, supporting the aim of my study in further understanding how evaluation is practised in a DE context. However, the quest for understanding the whole by looking at the particularities of a single case that can be transferrable is a key aspect of my study. Next, I explain why a

single in-depth case study was conducted, describing the embedded element of its design.

5.2.4.1 Single in-depth case study

This research is an exploratory investigation of the EP of a single project – the YP - led by the Northwest Development Education Centre (NDEC). Hence, I adopted a single in-depth case study approach to provide a “*thick description of the incidents in the everyday life*” of YPEP over time (Simons, 2009, p.3). Simon’s (2009) approach subscribed Yin’s (2008) definition of a single case study design as:

an empirical inquiry that investigates [a] contemporary phenomenon in depth and within its real-life context. This means that one single case study can entail all elements that needed to be included in this design: unit of analysis/social phenomenon, real-life context and no evident boundaries between context and phenomenon. (Yin, 2008, p.13)

The *in-depth* element means that a thick description is developed during the data analysis and interpretation to examine how the YP evaluation has been practised and changed during a three-year period. It does so, by exploring the project’s EP over time and how three groups of participants - coordinators, practitioners, and funders - experienced EP (see section 5.3.5; composite profile of the research participants). In this respect, I draw on Yin (2008) to investigate how evaluation has been practised from the perspective of different participants, illustrated as embedded units of analysis in Table 5.1 (overleaf):

Embedded units of analysis	NDEC	PDEC	DfID	GMA
Coordinators	1	1		
Practitioners	5	2		
Funders			1	2

Table 5.1: Embedded case study (source: Yin, 2008)

The embedded case design means that the main unit of analysis is the YPEP that is investigated through data collected from three embedded units of analysis (Yin, 2008, p.43). Two reasons support the rationale for this research design: firstly, the recognition that the participants' experiences could be collected and analysed into embedded units of analysis over time, ensuring the *in-depth* element of this exploration. Secondly, the separation of the embedded units allows the researcher to obtain valuable knowledge about the case as a whole (Yin, 2008).

Consequently, a single in-depth case study approach appears to be appropriate to accomplish the aims of this research. Nonetheless, two other sets of criteria reinforce this methodological decision. First, what I call 'researcher loyalty', which means that I ended up researching a problem that I closely experienced. In other words, I had seen this research problem occur while I was volunteering in NDEC, precisely when they had been asked to evaluate the YP, and were struggling to do so. Having considered six other DE organisations it turned out that NDEC, actually, contained all the relevant aspects necessary to conduct a single in-depth case study. These are:

- Social phenomenon to be investigated;
- Various embedded units of analysis within the case;
- Real-life context and no evident boundaries between context and phenomenon;
- Easy access and rapport;
- Accessible location allowed for data collection on a more regular basis (Simons, 2009)

Second, previous scholars have called for further research, particularly, in-depth qualitative studies that look at how NPOs evaluate their interventions in a DE setting (Carman, 2007; Bourn, 2014). In emphasising the particularisation of this design, the in-depth element of the case study allowed me to depict how EP has changed over time and under what circumstances. As such the specificity of focus makes a single case design adequate for exploring “*puzzling occurrences arising from everyday practices*” (Merriam, 1998, p.29). In summary, this thesis’ research design combines two distinctive features: first, Simons’ (2009) case study approach applied in other evaluation studies and, second, Yin’s (2008) feature of an embedded case study.

5.3 Approach to fieldwork

5.3.1 Situating the research and the researcher in context

Some scholars have positioned the role of researcher as central within a qualitative study (Denzin and Lincoln, 2005; Simons, 2009). The act of situating the research and developing an appropriate research methodology begins with a “*socially situated researcher*” (Denzin and Lincoln, 2005, p.8).

It is through analysis and interpretation of how people think, feel and act that many of the insights and understanding of the case are gained. It acknowledges that you are the main instrument in data gathering, interpretation and reporting. While this is significant in all forms of research, **in a single case and with qualitative methods, the “self” is more transparent and it is important to monitor its impact on the research process and outcomes.** (Simons, 2009, p.4, **emphasis added**)

In order to clarify the methodological choices adopted, it is important to place myself as a researcher within the social situation being explored. This research project was developed from my own experience as a practitioner in the field of evaluation and it was that hands-on experience that placed this research within EP (see section 1.2.1; the journey). In addition, the focus on the DE sector arose, partly, from my previous involvement as a volunteer with NDEC, when beginning to plan the evaluation of the YP, and partly from my past experience as a consultant with other NPOs. My involvement and observation of the developments at NDEC, as well as my close interest in the YP evaluation led to the research questions being structured to answer questions arising directly from practice and having a direct relevance to the EP of the case organisation, and ultimately to my own development as a practitioner, in the evaluation of social interventions. The subsequent refinement of the research questions was done in the light of phase two of the literature review.

As outlined above, to situate my role as a researcher, it is vital to recognise my previous roles as a volunteer and practitioner within the field of evaluation. In the past I have worked closely with the case organisation (NDEC), assisting mainly in the YP evaluation design and planning stage. I also had some previous partial information of the NDEC context, their financial constraints, their partnership agreements, the DE policy framework in which they operated and

the funder's demands for the YP evaluation. As expected, my previous professional experiences shaped the way I understand the world, so it is fair to acknowledge that as a practitioner I already have my own views on evaluation that have been challenged by my role as a researcher. This dilemma resonates with the idea that researchers always have to "*draw on their existing theoretical knowledge in order to understand, describe and explain social phenomena*" (Kelle *et al.*, 1995, p.12). Although I had an impression that some performance-based approaches to evaluation were inadequate, it was only through an in-depth exploration of a single project's EP over time that I observed the constraints that this approach seemed to perpetuate.

5.3.2 Social practice theoretical framework employed

The beginning of my research journey was heavily influenced by social practice literature, particularly applied in the evaluation work of Saunders (2000; 2005, 2011), Trowler (1996, 2013; 2012), and McCluskey (2011). The approach taken to fieldwork was influenced by the latest developments and further applications of social practice theory (SPT) in evaluation, as well as in other domains. The theoretical approach taken defines practice as "*a routinised daily task, manifested through a set of behaviours, which form ways of knowing how, desiring, wanting and avoiding*" (Reckwitz, 2002, p.250). Saunders, Trowler and Bamber (2011) draw on this definition to re-interpret that whatever one does in an evaluative setting can be seen as a practice, with several elements of mind, body, structure, language and knowledge. For example, the writing of a funding bid may be perceived as evaluation; as well as the daily conversation with a colleague on how to demonstrate the worth of a given intervention. For this

reason, this research is important to gain fresh insights about how evaluation is perceived and practised over time, within a DE context.

Previous scholarly attention focused on the potentialities of exploring practices (Bemelmans-Videc *et al.*, 1998; Kushner, 2000; Schwandt, 2002, 2005), which paved the way for the use of a SPT framework in novel research settings, such as DE. Moreover, previous scholars have noted the importance of interpreting evaluation as a social practice, because the depiction of practice is vital to understanding evaluative processes, particularly when they are outcome-based and concerned with social change (McCluskey, 2011; Saunders *et al.*, 2011; Trowler *et al.*, 2014) as it is DE. In this vein, a SPT view of evaluation considers practitioners and evaluators as constant learners on the art of evaluation, principally, on how to adapt, reflect and reconfigure their practice (Saunders, 2000; Saunders *et al.*, 2011). The framing and application of SPT in evaluation has informed my decision regarding its use in this research, mainly in four aspects:

1. Evaluation purpose – A SPT view of evaluation aims to depict learning; thus one assumption of my research is that the main purpose of evaluation is learning; whether in the form of recommendations to improve further interventions; or learning about the practice of evaluation itself (Saunders, 2000; Saunders *et al.*, 2011; Schwandt, 2015).

2. Timeline-based – This research explores how evaluation is practised and changed across a project timeline. Previous work on how to plan and conduct a

project evaluation considered the timeline of a project as a fundamental tool to depict its EP, as was the case of RUFDATA (Saunders, 2000).

3. Reflection – Previous SPT studies illustrated that a practice-based approach to evaluation values a reflexive process as an active involvement and participation of practitioners, alongside an awareness of the multiple meanings attached to EP (Saunders, 2000, 2011). My research investigates if and how practitioners perceive what they do as evaluation and, in turn, what do they perceive by evaluation.

4. Capture change – In considering the timeline of YPEP, a SPT view privileges a reflexive-based enquiry for planning the evaluation, capturing how EP has changed over time. In using a SPT lens, my research aims to depict how the YPEP has changed over a three-year period.

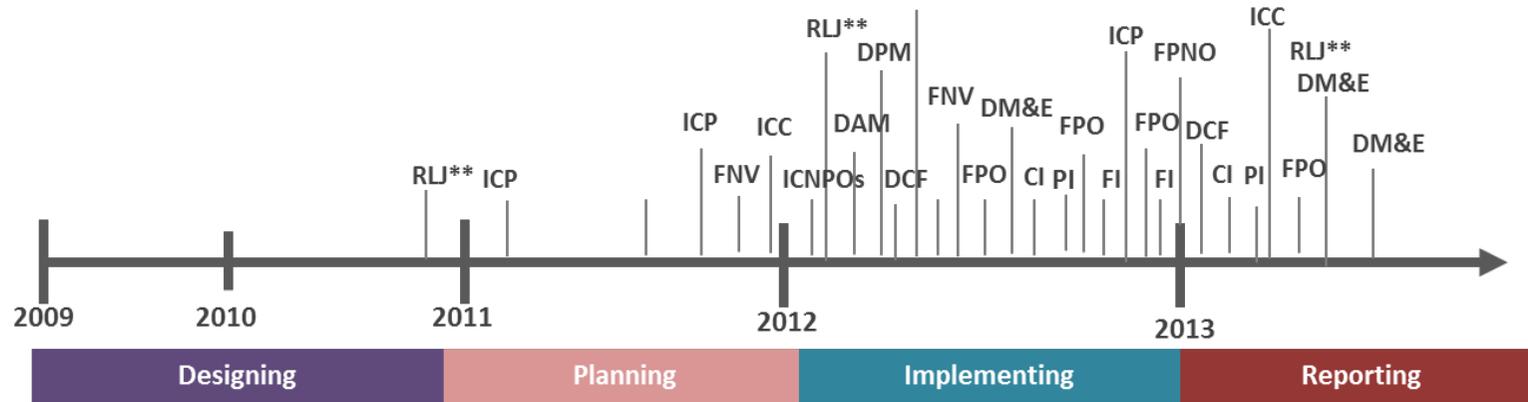
A final note is that, although, my initial approach to fieldwork has mainly been informed by Saunders, Trowler and Bamber's (2011) view of evaluation, as the research unfolds, I found a recent version of SPT, the dynamics of social practice (DSP) (Shove *et al.*, 2012), which surprisingly has not yet been substantially applied to research evaluation; therefore, DSP is the refined theoretical and analytical framework used in the thesis (see section 5.3.7; data analysis and interpretation).

5.3.2.1 Longitudinal dimension explained

Case studies include many variables and portray their interaction, often, over a period of time (Merriam, 1998), thus they can be longitudinal cases (Huber and Van de Ven, 1995). Given the impossibility to have followed the YP since its beginning, a retrospective research design (Denzin, 1989) enabled the collection of data in two phases, gathering materials from the three stages of the evaluation: planning, implementation and reporting of the evaluation.

A possible limitation of this retrospective design is the influence of participants' sharing of past events on the interpretation of data. Some may not have remembered important details or unintentionally obfuscated aspects that could have altered the analysis of data. To overcome this limitation, I relied on data triangulation and interviewed different participants about past events (see section 5.3.8.1). For instance, in one case, in which the flow of their story was unclear, I had informal conversations with that participant for further clarification.

Figures 5.4 and 5.5 (overleaf) illustrate the multiple points of data collection and the multiple sources of data covering each stage of evaluation. The longitudinal nature of data collection was critical to the aims of this research exploring how evaluation is practised across a three-year timeline, as well as how it changed over time. For the purpose of answering my research questions a longitudinal dimension was employed.



** Researcher’s learning journal used in retrospect during volunteer and pilot study period

Multiple sources of data and youth project’s stage covered

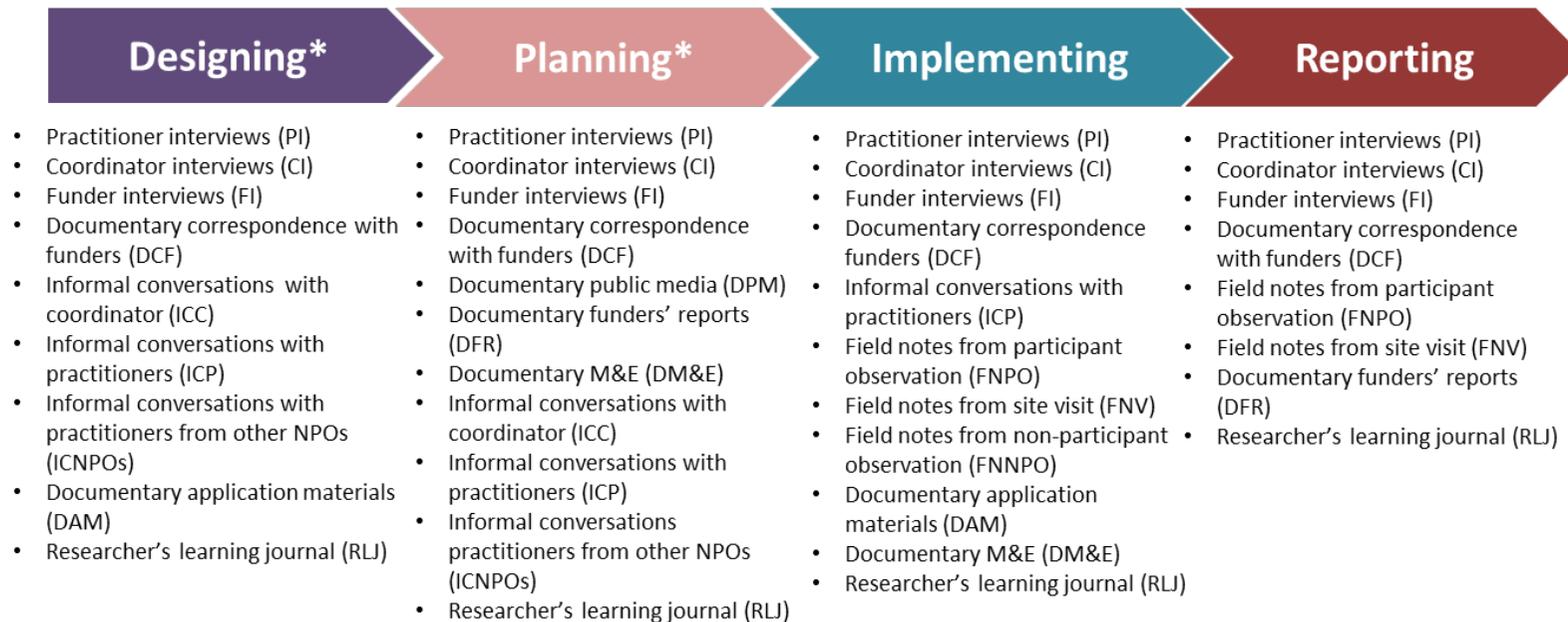


Figure 5.3: Multiple points and sources of data collection per stage

Data Source		Point(s) of data collection
CI	Coordinator interviews	Designing → Planning → Implementing → Reporting
DAM	Documentary application materials	Designing → Planning → Implementing → Reporting
DCF	Documentary correspondence with funders	Designing → Planning → Implementing → Reporting
DFR	Documentary funders' reports	Designing → Planning → Implementing → Reporting
DM&E	Documentary M&E	Designing → Planning → Implementing → Reporting
DPM	Documentary public media	Designing → Planning → Implementing → Reporting
FI	Funder interviews	Designing → Planning → Implementing → Reporting
FNNPO	Field notes from non-participant observation	Designing → Planning → Implementing → Reporting
FNPO	Field notes from participant observation	Designing → Planning → Implementing → Reporting
FNV	Field notes from site visit	Designing → Planning → Implementing → Reporting
ICC	Informal conversations coordinator	Designing → Planning → Implementing → Reporting
ICNP	Informal conversations with practitioners from other NPOs	Designing → Planning → Implementing → Reporting
ICP	Informal conversation with practitioners	Designing → Planning → Implementing → Reporting
PI	Practitioners interviews	Designing → Planning → Implementing → Reporting
RLJ	Researcher's learning journal	Designing → Planning → Implementing → Reporting

Figure 5.4: Multiple sources of data per alphabetic order

5.3.2.2 Ethnographic influence explained

Literature on ethnographic studies on evaluation affirms that one of its most important contributions is “*its ability to provide the context required to interpret data meaningfully*” (Kellaghan and Stufflebeam, 2003, p.51). For this reason, an ethnographic component of this research is fundamental in providing the evaluation’s contextual information. That was achieved through my immersion in the field, as a native, submersing myself in this particular setting for a period of one year (Silverman, 2011). The ethnographic element in this research was prominent during data collection, where I gained access to what people actually do, in contrast to what they might state they do (Mintzberg, 1980). Additionally, I had the possibility of not simply observing what participants’ realities were, but also sensing it, through becoming one of them at certain times (Denzin and Lincoln, 2005).

My adoption of an ethnographic stance to this research was also informed by Trowler’s (1996) design of a quasi-ethnographic single case, where he supported the combination of the single in-depth case study with an ethnographic data collection approach. Trowler’s account was helpful to illustrate how an ethnographic single case has been written up, in terms of structure and style for an in-depth descriptive account. Besides, it was the choice to collect data ethnographically that reinforced my identity as an observer, which had to be acknowledged and monitored throughout the research journey (Simons, 2009). Lastly, an ethnographic element was essential to pursue the conduct of a single embedded case, in which data has to be collected from three sources: coordinators, practitioners and funders.

5.3.3 Research setting

The formulated research questions shed light on the need to gain an in-depth understanding of the research participants' context. In this respect, a number of challenges were faced, for example, the need to establish a close relationship with participants to allow an ethnographic exploration of the evaluation; the understanding of how the wider context influences EP in a DE domain; the refinement of my approach to fieldwork to investigate EP without disclosing a social practice view as important, that can potentially bias; and the guarantee that a longer interaction with the research setting was in place to allow a longitudinal observation of the setting. The latter was particularly challenging for a part-time student.

Past research on EP confirms that research settings that allow for longer interaction with participants generate a rich dataset (Saunders, 2000; Robinson and Cousins, 2004; Ding, 2009), which I had already anticipated from my previous experience as a volunteer with the case organisation (see section 1.2.1; the journey). I also knew that NDEC was running a three-year funded project, and was about to plan, implement and report its evaluation. Based on these insights, the YP evaluation was selected as the main research setting.

The YP partnership has one lead-coordinator and five project workers, who were allocated a reduced number of hours. It has one volunteer who delivered diverse support and an external evaluator who was appointed on behalf of the funder's scheme, the Development Awareness Fund (DAF). The YP was funded by DfID, which had requested a mandatory evaluation to be reported at the end of the project. Furthermore, DfID sub-contracted a 'Grant Management Agency'

(GMA) to monitor and evaluate the projects funded under the DAF scheme, and as result, an external evaluator was selected to support the YP evaluation over time. In addition to the three groups of participants included in this research, others were purposefully excluded because they had not participated in the evaluation (the NDEC trustees and beneficiaries).

Having assumed a premise that participants' perceptions and attitudes are socially constructed, it was consistent to also examine their contextual environment, exploring NDEC as a case organisation. As a research setting includes the situated and the wider context in which the phenomenon occurs (Patton, 2011), the context of NDEC as a leading organisation matters; as well as the considerable fluidity around the conceptual definition of DE. Whilst some have conceptualised DE as "*foster[ing] the full participation of all citizens in world-wide poverty eradication, and the fight against exclusion*" (DARE Forum, 2004); others have debated the multiple ways of achieving this aim, of which global youth learning methodologies is only one of many approaches (Bourn, 2008; Adams, 2010). See Appendix 2.a; background information on DE.

Summing up, the main reasons leading to the selection of the NDEC -YP - as a research setting to explore EP are:

- A DE project, as a research setting allowed for the exploration of participants' experiences of EP.

- The design and planning of the YP evaluation permitted a full application of SPT methodologies to investigate how evaluation has been practised over time.
- YP was being implemented when I start volunteering in NDEC, so an atmosphere of trust was already established which made an easy access and rapport, enabling the conditions for an ethnographic immersion in the setting.
- Lastly, the parallel reading of DE literature accounted for the adequacy of this research setting, as most recommendations called for further research in loco, with real-life time events through the depiction of a variety of stakeholders' views over time.

5.3.4 Sampling procedures

NDEC as an organisation was selected as an *information-rich case* study (Patton, 2002, p.230) from a sample of six other DE NPOs because it responded to the established sampling criteria: the level of DE activities within a project or program, in receipt of public or (private) funding; formal evaluation required by the funding agency; and located within the Northwest of England. From this rich case, the selection of participants was based on purposeful sampling, as their experiences were critical for an in-depth qualitative enquiry (ibid.), thus all twelve individuals who directly or indirectly participated in the project's evaluation were selected as primary research subjects. The purpose of purposeful sampling is to select information-rich cases whose study will illuminate the questions under study (ibid).

However, some groups appeared to have not been involved with the evaluation (trustee members and project's beneficiaries), so for purposes of anonymity, research subjects were allocated to one of three participant groups: coordinators, practitioners and funders.

A possible limitation may arise from the fact that I already had contact with the case organisation. In acknowledging this limitation, I believe that I used the correct tools to monitor my own bias, in the awareness that this case would generate fresh insights about EP. As unpredictable things happen, I met two other potential organisations within the initial sample, just in case the selected organisation failed. Also, as the data collection period was approaching, I devised a composite profile for my research participants.

5.3.5 A composite profile of the research participants

This section describes three composite profiles, using gender neutral names to ensure the confidentiality of the information shared by my research participants and to protect their anonymity (Greene, 1988). The purpose of these profiles is to allow a depiction of each group's roles in terms of their participation in the evaluation process and in which capacity they engaged with the YP evaluation. The profiles of coordinators, practitioners, and funders are visualised in Tables 5.2, 5.3 and 5.4, respectively; from which I describe some general features after each profile.

Coordinators

Glenn and Eli Typical DE coordinator profile

Background

Eli holds a degree in education and development studies, but no specialisation in evaluation. Having very little experience of evaluation and sometimes negative past experiences alongside their professional journey, they heavily rely on project managers to monitor and evaluate projects. Eli “*learns by doing it*” and by internalising the project cycle management pattern that works for funders.

Main working activities

Eli has to keep the organisation running smoothly, in terms of management, funding and human resources and is ultimately responsible for implementing, evaluating and reporting funded projects to the funders and trustees. When working in partnerships, Eli has to make sure that the team delivers accurate reports on time.

Other working activities

Eli’s other activities are organisational management and strategic direction of the organisation. Decision-making is also part of the daily routine, as Eli has to provide funding and plan the organisation’s financial sustainability.

Main challenges

With the austerity measures, Eli has to deal with rapid and unpredictable change in the DE domain. As a result, this has improved Eli’s creativity in finding solutions for current challenges.

Table 5.2: Composite profile of a DE coordinator

Within my research participants, I interviewed two coordinators, one from the United Kingdom and the other from overseas. In terms of their academic background, both held a postgraduate degree - one held an MA and the other was doing a PhD. Their professional experiences were from activism and

teaching through to community-based work with other charities, social cohesion programs and consultancies within the DE sector. Both coordinators had already worked with other DE centres (DECs). Neither of them reported a professional expertise in evaluation even though one had been as a program evaluation consultant in their last job abroad.

For the purpose of the thesis, I refer to the group as ‘coordinators’ and I mention the respective gender-neutral pseudonym, when referring to a specific participant. On a few occasions I had to stick with the abbreviation ‘DEC’ and withheld participants’ composite profile and the organisation’s pseudonym, because the topic risked the disclosure of participants’ identities. This extra measure ensured that their anonymity and confidentiality was protected.

Practitioners

**Addison; Dale;
Jesse; Julnes;
Sam; Stevie and
Tyler**

Typical DE practitioner's profile

Background

Julnes' background is in theatre studies, working with young people in arts-based performances. Julnes has little experience in evaluation, despite previous participation in some evaluations where the tasks were prescribed by the coordinators. They practise evaluation in a mechanical pattern, so their perception is that evaluation is not for everybody. According to Julnes, there are specific skills one has to have, and this is why external evaluators, quite often, support DE evaluations.

Main working activities

Generally, Julnes does not have a main practice allocated and has to do whatever is needed. Although their contract is assigned to a specific project, in reality they work for various projects simultaneously. They monitor and evaluate the project's implementation but find it hard to measure its performance in order to demonstrate the evidence achieved to the funders. Julnes is contracted on a weekly number of hours-basis, which makes it difficult to complete all tasks assigned. For this reason, the organisation has volunteers to fill this gap.

Other working activities

Alongside the M&E, Julnes has to design other projects to receive funding grants, and tends to proactively search for alternative funding, which the coordinator often appreciates. Whilst writing these funding grants, they also plan, conduct and report other project evaluations.

Main challenges

Julnes struggles to translate the funders' vocabulary into meaningful requirements, for example, the meaning of outcome evaluation matrix, logic frameworks and theories of change.

Table 5.3: Composite profile of a DE practitioner

Seven practitioners were interviewed, mostly project workers with different contract hours; as well as a finance officer and a volunteer. All of them were from the UK, with only one non-UK national. In terms of their academic background, two participants held a teaching qualification whereas others had qualifications in finance, youth work and technology (IT). Their professional experience included various temporary occupations, such as teaching, activism, finance, administration, IT, fundraising, and youth work. For the purpose of the thesis, I refer to 'practitioners' when broadly mentioning to this group of participants and to the respective gender-neutral pseudonym when stating a specific participant.

Funders

**Charlie, Kelly
and Brook**

Typical DE funder

Background

Charlie holds a degree in Economics plus a specialisation in development studies, and used to work as evaluation officer in funding agencies, although their first evaluative experience was overseas as a volunteer. One of their areas of expertise is to make clear analysis of complicated situations, problem solving, and management of complex issues.

**Main working
activities**

Charlie is responsible for keeping a strict control and to “*check out*” whether the standards of performance are met by grant-holders. For example, Charlie’s main activities include responding to grants’ finance queries, processing payment claims, reviewing project planning documents and project progress reports. Additionally, they also support organisations thinking about future strategy towards sustainability.

**Other
working
activities**

Charlie has to externally evaluate some grant-holders interventions, to assess applications for funding approval, as well as to manage other funding schemes. That is especially problematic near deadlines for submission, when grant-holders request more support and clarification.

**Main
challenges**

Charlie provides support to grant-holders about reception document plans and points to the respective explanations in their M&E handbook. As a funder, Charlie holds expectations about grant-holders’ performance and for that other projects’ good practices are shared with them.

Table 5.4: Composite profile of a DE funder

The funders included one DfID employee, one GMA program manager and one self-employed external evaluator, who was also a GMA fund manager. They were all from the UK, and two of them had worked abroad before. Regarding their academic background, they held degrees in management and in education and were termed as generalist workers. One of the funders' participants did not have any specialisation in development studies, however, in terms of their professional experience all participants had worked with charities before and one of them had supported strategic processes within the DE sector. For the purpose of the thesis, I refer to the participants within this group as 'funders' and to the respective gender-neutral pseudonym when mentioning a specific participant.

The decision to group the research participants into three embedded cases, according to their professional role, ensured that their anonymity and confidentiality was protected. However, I realised that I could not have exposed the external evaluators' profile in the same profile as above, because there was only one person interviewed in that capacity whose articulations appeared influential to the exploration of the EP. As a result, I composed these typical profiles drawing particularities from multiple persons, funders, and the external evaluator to ensure that all my research participants' anonymity and confidentiality would be protected.

In summary, these composite profiles draw on information from various individuals and for that reason the information was combined into broader profiles. Also, there are common and uncommon features that characterise each

group of participants, thus these descriptions allow for a broader representation of participants.

5.3.6 Data collection methods

The section starts with the data collection approach used in this case study, to then describe the methods used and why. The section briefly describes the first part of the research in practice, through a pilot conducted. Table 5.5 (overleaf) illustrates the preferred methods in case study research

Research questions	Data collection questions	Preferred methods	Epistemology	Primary values	Key audience
Overarching RQ: How do DE organisations evaluate their social interventions?	How do participants describe their tasks when evaluating the YP)				
RQ 1. What are the NDEC's stakeholders doing when they are evaluating the YP?	How do participants describe their tasks when evaluating the YP? What do practitioners do at the same time when evaluating the YP?				
RQ 2. What are the influences that shape the NDEC's evaluation practice, particularly the YP?	How was the YP designed? What was the external evaluator's main role in the context of YP's EP? To what extent does the context influence the NDEC's EP? Give an example, please	Qualitative: case study, semi-structured interviews, observations, document reviews	Social Constructionism	Pluralism; understanding personal experiences; observing practices	DE practitioners, project coordinators, funding agencies and NPOs community
RQ 3. How has evaluation practice changed across the YP's evaluation timeline?	Would you walk me through other NDEC's previous evaluations? What were the main lessons learned? How similar/different has the YP evaluation been?				

Table 5.5: Preferred methods in case study research (adopted from Simons, 2009)

Two initial interviews and informal conversations were conducted as part of a pilot study with practitioners working in other NPOs in the UK, during February 2012. The pilot exposed the complexity of interviewing in English as my second language, and made explicit the profundity of practitioners' concerns about EP. This notion reinforced my decision to conduct an in-depth single case study, and to start data collection at an early stage, as time was needed to be immersed in the field, as well as receptive to the sharing of my participants.

Two specific refinements were made after the pilot research: first, a selection from all antecedents of EP reviewed in the literature (Smith and Brandon, 2008). After these interviews, it was clearer that the three antecedents that informed practitioners' evaluation the most were power, context, and perception of evaluation; so these were likely to apply to the subsequent research participants. Second, in refining the interview protocol, I purposively omitted the term 'practice' and when possible the term 'evaluation', to allow freedom for participants to articulate their visions, particularly in seeking to understand what do they mean by evaluation itself. This was challenging and not always possible, however, I monitored the conduct of the interviews in relation to the latter through an interview debrief form (see Appendix 4.d).

5.3.6.1 Primary data sources: semi-structured interviews, observation, and documents

Three sources of data followed from the application of case study research procedures: semi-structured interviews, observations and documentary analysis. Data from each participant was collected in two phases over time (Denzin and

Lincoln, 2005; Simons, 2009), covering the three stages of the project evaluation: planning, implementation and reporting. The data collection period described next is organised by the full explanation of each primary data source followed by a table summarising the data collected.

- Semi-structured interviews

Figure 5.5 below represents a broad overview of the two phases of data collection, alongside my presence in the field.

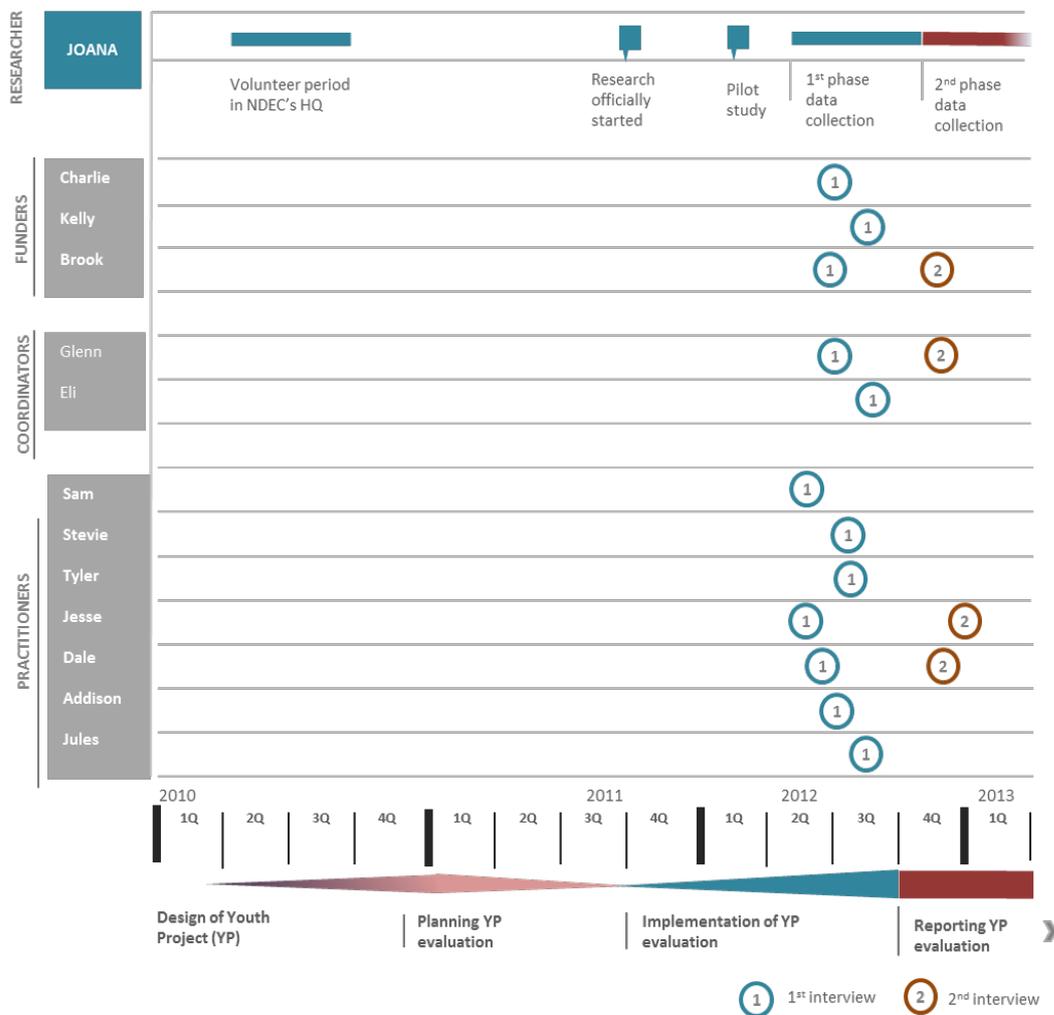


Figure 5.5: Phases of data collection

Participants were asked to describe their involvement in the YP since the beginning, as well as their current roles and responsibilities. They were also asked to outline the tasks they usually do, and associate with evaluation; through filling a form with evaluation activities (see Appendix 5; evaluation form). I conducted sixteen semi-structured interviews over a period of one year (March 2012 until March 2013) with individuals in three groups of participants, who dealt with the YP evaluation on a regular basis. Twelve semi-structured interviews, lasting one and a half hours, were conducted in phase one and four interviews were conducted in phase two, as illustrated in Table 5.6, below:

Phase	Date	Interviews
1	March-Dec 2012	12
2	Jan-March 2013	4
Total		16

Table 5.6: Semi-structured interviews

This research adopted a responsive interviewing approach for four reasons: first, it requires the development of a trusting personal relationship between the “*conversational partners*”(Rubin and Rubin, 2012, p.7) and that fitted with my ontological and epistemological views, because it advocates that both researcher and participants are socially constructing one interpretation of a social phenomenon through experiencing a “*joint process of discovery*” (ibid. p.36). Second, a responsive interview approach advocates a certain level of flexibility concerning the research design; and third, it adapts to a variety of interviewing situations. As an example, I had planned to interview two participants within the

funder's group – the external evaluator and the strategic advisor - but then realised that these roles were assigned to the same participant. The responsive approach allowed me to use the same interview protocol, to focus on these two different roles, thus prompting the same participant from two different angles (see Appendices 4.b and 4.c; first and second interview protocols). Fourth, the responsive interview combines well with other qualitative methods, especially participant observation. Overall, this approach assumed that what participants have experienced is true for them and, by sharing it they allow my entrance into their world.

As some of the events regarding EP occurred in the past, and were influenced by participants' previous experiences, sometimes a direct observation was not possible. Thus, a semi-structured interview gathered data regarding meanings, values and context of these past events to participants' current EP (Patton, 2002). Responsive interviewing also focused on forming a mutual relationship with each participant that often outlasted the period of research. As this relationship was based on "*trust, reciprocity, and acceptance*" (Rubin and Rubin, 2012, p.37), the responsive approach was consistent with my previous volunteering experience in NDEC, through which I knew some of the participants. To a certain extent, this previous contact with my "*conversational partners*" (ibid. p.7) proved useful regarding access, negotiation and rapport during the fieldwork. The core of a responsive interview involves the formulation of three types of questions: main questions, probes, and follow up questions (ibid. 2012), as Table 5.7 (overleaf) illustrates. The main questions

addressed the overall research and set the structure of the first phase of interviews, and within this question, the participants were asked to describe in detail the stages of planning and implementation of the YP evaluation.

Main question	Could you walk me through your role within the YP, since its onset?
Probe question	Which of the following activities have you used during the planning, implementation, and reporting stage of the evaluation?
Follow-up question	Regarding the purpose of evaluation you have said something like... <i>the results were on paper, so we can meet the targets for the funders to be pleased with that...</i> What do you mean by “ <i>please the funders</i> ”? In your opinion, what is the funder’s main purpose of evaluation?

Table 5.7: Responsive interview type of questions (Rubin and Rubin, 2012)

The probes helped to manage the conversation as well as clarify some aspects of the participants’ responses. For example, the participants received an evaluation form with various evaluation activities and were asked to expand on those they recognised as evaluative. In this regard, participants were not prompted about their practices, but asked to describe routinised tasks related or included in the YP evaluation. Through the probe questions, participants had the opportunity to develop their reflection and many times they asked for additional time to think, before articulating their answers. The follow-up questions explored ideas that had emerged during the first interview that needed further clarification and these were extensively used in the second phase of data collection. As a result, the

second interview consisted of mostly follow-up questions, clarifying previous issues or asking participants to further elaborate some specific aspects.

In the second phase of data collection, I conducted four semi-structured interviews with participants, who were most involved in the evaluation. For the purpose of the thesis, the four interviews are referred to as ‘2nd interview’ when directly quoted, and the remaining interviews are referred to with the participants’ pseudonym. It is relevant to acknowledge that I did try to arrange a second interview with the other participants, but they were unavailable during that period. The second interviews were one hour in length and based on the same responsive approach as the previous (Rubin and Rubin, 2012). They were robust in the sense that both, participants and I, were fully immersed in the content, therefore the flow and reflection was greater than the first phase of interviews.

In summary, the strength of the responsive interviewing approach appeared in the “*researcher's ability to hear what was said; to change direction to catch a new insight; to track down a new theme; or to refocus the broader questions*” (ibid. p.42). However, ethically speaking, there are some considerations that need to be taken into account, as explained in section 5.3.9; ethical considerations in the fieldwork.

- Observation

To capture data through a combination of data sources appeared crucial to “*allow access to the observations of others*” (Weiss, 1995, p.1). Hence,

alongside the sixteen semi-structured interviews, a series of ninety-four hours of participant observation and twenty-four hours of non-participant observation complemented other sources of data, as Table 5.8 (overleaf) illustrates:

Phase	Date	Participant Observation (hours)	Non-participant Observation (hours)
1	March-Dec 2012	71	18
2	Jan-March 2013	23	6
Total	118	94	24

Table 5.8: Observation sessions

The lived observation of the participants' EP was the fundamental basis for my deep reflection throughout the research journey. For the purpose of exploring how the project was evaluated, I observed various instances involving the three groups of participants, which meant the main location was the NDEC office, with some visits to other sites, official events, and extraordinary meetings. Regarding the funders' group, all interviews were conducted and recorded through Skype.

During the first phase, I also conducted eighteen hours of non-participant observation in some events and meetings, included in the daily management of the project. During the second phase of data collection, I conducted twenty-three hours of participant observation, mainly in the NDEC office, and six hours of non-participant observation in an extraordinary meeting, during the reporting stage of evaluation (see Figure 5.3; multiple points and sources of data collection per stage). The incremental focus of the observations was gained due

to a longer immersion in the field, which allowed me to capture important details, for example about the communication between the three groups of participants, as well as the main differences in their narratives, tone of discourse and perception of evaluation. These were captured through field notes from the observational sessions, organised and coded in ATLAS ti.

- Documentary evidence

The collection of key documents gathered relevant background and official information about the YP. A total of thirty documents were analysed after an initial examination of many others, such as DE policy reports, NPOs publications, and internal archives from the YP needs assessment.

Phase	Date	Documents
1	March-Dec 2012	27
2	Jan-March 2013	3
Total		30

Table 5.9: Documentary evidence

The main documents analysed were assigned to one of five categories to support the organisation of data, in ATLAS.ti, and to subsequently contrast with other sources of data. Table 5.10, below, shows the categories and number of documents analysed across the two stages of data collection.

Categories of documents analysed	Doc. application materials (DAM)	Doc. Correspondence with funders (DCF)	Doc. funder's reports (DFR)	Doc. M&E reports (DME)	Doc. public media (DPM)
Phase 1	5	10	2	7	3

Phase 2	-	-	3	-	-
Total: 30	5	10	5	7	3

Table 5.10: Categorisation of documentary evidence

The DAM category includes the YP proposal and four technical and financial forms. The DCF category includes four formal letters and seven emails. The DFR includes three intermediate and final annual progress reports, one project completion report and one external evaluator’s report. The DME has two monitoring and evaluation documents, one logic framework, one evaluation outcomes matrix and three evaluation activities sheets. The DPM category includes two videos, and one press release. The documentary evidence analysed in the second phase was from the project’s final completion report, the third-year intermediate report and the external evaluator’s report, because these were only finalised at that time.

In summary, the overview of the two phases of data collection is visualised in Table 5.11 below:

Phase	Date	Interviews	Participant Observation (hours)	Non-participant Observation (hours)	Documents
1	March-Dec 2012	12	71	18	27
2	Jan-March 2013	4	23	6	3
Total		16	94	24	30

Table 5.11: Overview of the data collection phases per data sources

5.3.6.2 Secondary data sources: informal conversations and learning journal

Other sources of data included extracts captured in the form of a learning journal to allow further reflection (Moon, 2006), and informal conversations (see Table 5.12, overleaf), which were held twice with four participants (Patton, 2002), in a relaxed environment that allowed participants to share relevant data in a different way. These conversations were spread across the data collection period with coordinator and practitioners, and personal observations were captured in my learning journal. As a result, I found that there were too many fascinating issues to uncover, some more related to perception, competence and overall purpose of evaluation, whereas others related to funders' power and compulsory requirements when reporting evaluation.

Phase	Date	Informal conversations	Extracts
1	March-Dec 2012	4	
2	Jan-March 2013	4	
Total		8	1 learning journal

Table 5.12: Informal conversations and extracts collected

5.3.7 Data analysis and interpretation

It does not matter what my paintings will be worth in a hundred years' time or that people will see what I see. What matters is that they see differently. (Magritte, 1979 cited in Stake and Kerr, 1995, p.56)

This section describes the approaches that influenced the process of data analysis and interpretation, and details the thematic analysis undertaken and the

respective protocols employed. The analysis of data was underpinned by the influence of various approaches, as illustrated in Table 5.13 (overleaf):

Approaches to data analysis and interpretation	Main focus	What I take to my analysis?	Why?
Retrospective analysis (Saldana, 2002)	Characterised past events as ‘epiphanies’ exploring participants’ views.	Retrospective element in the interpretation of past episodes that may have generated change. (Practice-change episodes (PCEs)).	Given the impossibility to follow the YP since its design, some events were retrospectively analysed according to how they affected the participants’ sense of EP at the time of data collection.
Iterative analysis (Huberman and Miles, 2000).	A succession of questions and answers to the data allowed the inductive nature of the research.	Inductive element of analysis	Because no pre-defined hypothesis had been set, I had various interactions with data during the whole period of analysis.
Transformation of data (Wolcott, 1994; Simons, 2009)	Make sense of the research materials, by acknowledging that a combination of analysis and interpretation, transforms data.	Notion that interpretation and analysis “ <i>dance</i> ” together, side by side, but sometimes one goes faster than the other (Simons, 2009, p.140).	In allowing the understanding derived from a “ <i>more holistic, intuitive grasp of the data</i> ” (Simons, 2009, p.117), it reveals in-depth insights appropriate for a thick descriptive account of this research’s single case.
Computer-assisted analysis Friese, 2012	Systematic exploration of data, through assisted preparation, coding, categorisation and query across the whole dataset.	NCT motto - noticing things, collecting things and thinking about things.	Allows a consistent analysis of a variety of research materials, accepts the non-linear nature of research by moving back and forth between noticing, collecting

and thinking about things.

DSP (Shove et al. 2012)	Framework to explore instances of change in EP over time.	Timeline of EP	<p>DSP enables the exploration of how EP changes over time by longitudinally observing how its elements –material, competence and meaning – are enacted and reconfigured.</p> <p>It examines the persistence and disappearance of practices, in terms of the changing relationships between their constitutive elements.</p> <p>A DSP lens has not yet been applied to empirical studies on EP, so it enables the understanding emerging data through theory.</p>
--------------------------------	---	----------------	---

Table 5.13: Influential approaches to my analysis and interpretation (source: author)

5.3.7.1 Thematic analysis step by step

This thematic data analysis was present in four steps, each one with a specific protocol:

1. Transcription and immersion into the data
2. Descriptive and process coding
3. Query data in ATLAS.ti
4. Reflective analysis refinement: a timeline of EP

Step 1: Transcription and immersion into the data

The first step to data analysis was the full verbatim transcription of the sixteen semi-structured interviews, supported by Dragon software. This period was critical to immerse myself in the data and, as a result, I started to make sense of the learning journal created during the data collection period. A first grasp of my data emerged in a messy and inconsistent flow, but with the support of a DSP analytical lens, I interpreted practice-change episodes (PCEs) to describe an instance of my account that emerged as influential to the project's EP and consciously or unconsciously reported by participants. Additionally, the thick description of this account followed the chronology of the project timeline, through data depicted from various dimensions of the case's real life, its political and organisational context, and participants' EP and working activities. As an original way to interpret participants' articulations, these PCEs represented their daily practices and it was through their descriptions and my interpretation of these episodes that change was captured across the YPET.

Step 2: Descriptive and process coding

In preparing data to be analysed, two challenges caught me by surprise. The first was the preparation of files to be uploaded in ATLAS.ti., as primary documents, as visualised in Table 5.14 below:

Original name_Dropbox_phd_fieldwork	Analytical name_dropbox_phd_analysis_Atlas Project set up
	UA1_2_4_5_6_7_Reports requested from donor
DAF Y2 Report	Dev Ed_donor_Y2 report_201106
DAF M&E long draft plan_201206	Dev Ed_donor_M&E long draft plan_201206
In-year_progress_report_Jan_2012	Dev Ed_donor_Interm progress report_201201
Matrix of learning outcomes	Dev Ed_donor_GMA adviser_Matrix Learning Out_201208
FINALY3 DAF PCR 2013	Dev Ed_donor_Y3 final Proj Completion report_2013
DAF 267 External evaluation Impact report	Ext Eval_donor_Ext Eval Impact report_2013
	UA2_Dev Ed int doc_Eval docs_Staff_HR docs
1. Mission, Values & Objectives	Dev Ed_Int doc_Mission statement
20121029 Project prioritising matrix_follow up_strategical meeting	Dev Ed_Strategical doc_Project Prioritising matrix_20121029
Evaluation report things to do checklist	Dev Ed_Int doc_Eval report check list_2011
M&E plan 2010-13 revised NOV_2011	Dev Ed_donor_M&E plan_revised_201111
Evaluation tools and reporting plan	Dev Ed_Int doc_Eval tools_reporting plan_2011
M&E Meeting Minutes	Dev Ed_Int doc_M&E meeting minutes_201104
M&E notes	Dev Ed_Int doc_M&E notes_2011
Outcome 3 Evaluation 2012	Dev Ed_Int doc_Eval form_Outcome 3_2012
July 2012 - evaluation data	Dev Ed_Int doc_Eval data_201207
Summary of Young people Outcomes 1&2	Dev Ed_Int doc_Eval data_young people outcomes 1_2_2012
Quotes	Dev Ed_Int doc_Eval data quotes_2012
student evaluation L & M college	Dev Ed_Int doc_Eval form_2012
teacher evaluations	Dev Ed_Int doc_Eval data_teachers outcomes_201207
	UA3 Dev Ed_Partner Docs
Partner Y3 M+E doc-1	Dev Ed partner_Y3 M&E doc_2013
	UA6_donor_think tank reports
DFID enablingeffectivesupport	Donor_enablingeffectivesupport report
Impact_of_global_learning_research_paper_web	Donor_think tank research paper_Impact_of_global_learning_2010
NWGEN Final Report 2010	Donor_NWdonor Edu network_Final Eval Report_2010
	UA7_donor_GMA docs
DAF Project Visits Jan-Feb 2012 - information for projects FNL-1	Dev Ed_Int doc_donor visit report_201201
Response to GMAs Y1 Report feedback	Dev Ed_Int doc_response to GMA Y1 report feedback_2011

Table 5.14: Preparing analytical data in ATLAS ti.

The process of transforming raw data into an organised and anonymised data set allowed the progress into the analytical categorisation, which was done through a list of primary documents (PD in ATLAS ti. terminology) uploaded to be analysed (see Appendix 7a; list of PD in ATLAS ti.). To overcome this challenge of preparing the dataset, I decided to rename each file according to the current research design status, and associate each participant with one of the three embedded cases, which was easily done through the ‘family manager’ function in ATLAS ti. This analysis was based on an inductive approach of

coding, which means the process of breaking down data into segments and assigning it a code (Simons, 2009). These codes were then ordered and examined for connections, patterns, and propositions, as illustrated in the network view in Figure 5.6 (overleaf). This was an initial messy attempt at descriptive coding, which was refined through various iterations of analytical coding, memos writing and querying data.

The second challenge was the realisation of the difference between descriptive and analytical codes, as illustrated below:

The screenshot shows the ATLAS.ti Code Manager interface with a list of codes. Two callouts provide definitions:

- Descriptive codes** are the main categories (e.g. context, evaluation process, evaluation perception).
- Analytical codes** are a refinement of specific associations, meanings, and patterns to make sense of data (e.g. evaluation perception_scepticism).

Table 5.15: Example of analytical coding in ATLAS.ti.

I then aggregated categories that seek to explain the data, through the production of themes (Simons, 2009) and as exemplified above, themes such as power relationships, evaluation process, evaluation purpose, evaluation practices, and evaluation perception appeared as relevant in a preliminary phase. Moreover, the richness of theming the data is that it is unlikely to be the only method used when immersing in the data (Saldaña, 2009). The combination of descriptive and process coding, during this first cycle of coding “*winnowed down*” the themes, to shed light on the essential ones (ibid. p.32). This description allowed me to note connections, curiosities, and possible patterns of themes across the YPET. This longitudinal element of the case study captured relevant changes in

participants' EP across the evaluation timeline, and so process coding¹² was appropriate to analyse “*ongoing action, interaction, emotion taken in response to situations or problems*” (ibid. p. 89).

Step 3: Query data in ATLAS ti.

A second cycle of coding paved the way towards answering the research questions. In the previous first cycle of coding, eighty-seven codes were generated, despite abstraction and lack of categorisation. As a result, a saturation point was reached, which meant that no new codes were added. During this cycle, I developed a more structured coding list through the classification, prioritisation and conceptualisation of the initial coding list, as illustrated in Figure 5.7 (overleaf).

¹² In process coding the aim is to capture verbs in the continuous form across data.

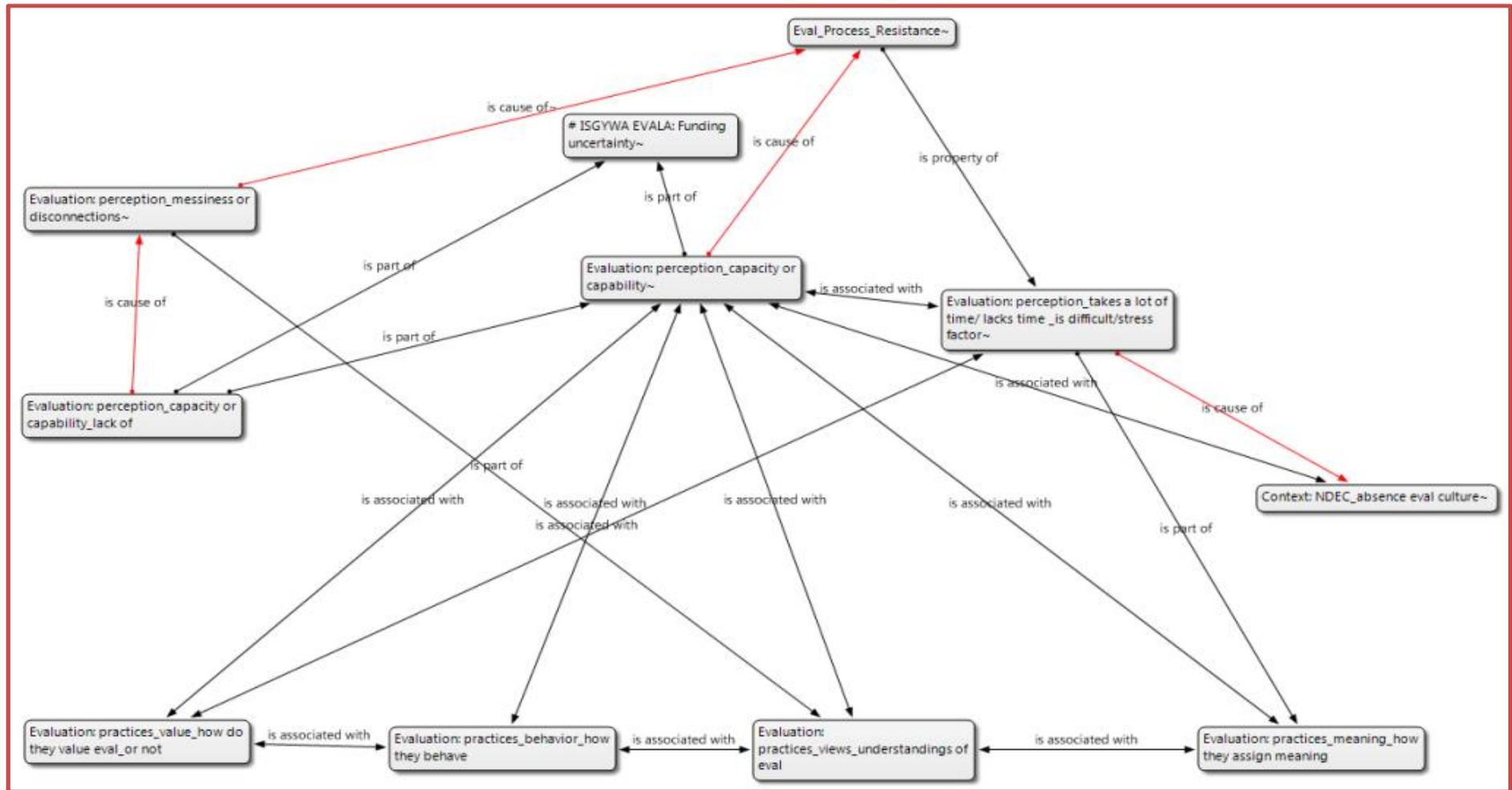


Figure 5.7: Network view refinement in ATLAS ti.

During the first cycle of coding, I analysed evaluation practice and evaluation perception as two separate codes to then realise that the former may be represented as, and influenced by a manifestation of participants' perception of evaluation. Thus, I queried both codes to find out that the way participants perceive evaluation seems to influence their practice of it. As this refinement seemed helpful to make sense of data, I applied the same principle with other codes, for example combining various themes through which power was manifested, such as context, relationships, and consequences.

The query tool in ATLAS.ti has a retrieval function that made it possible to combine more or less complex searches across the data set. For example, by combining the code 'evaluation practice prior GMA' with the family 'coordinators', I analysed specific segments per group, as well as refined preliminary findings emerging. I also queried data to analyse patterns of occurrence of preponderant themes in relation to the stages of evaluation – planning, implementation and reporting. For instance, a pre-initial stage of the project design appeared as influential to the subsequent practice of evaluation, so it was added to the final refinement of analysis as an unexpected finding. The stages of evaluation and its emerging themes were represented in a timeline of the project EP, which was then used to discuss the findings per theme. The final refinement of data analysis and the use of the timeline as an interpretative tool are described next.

Step 4: Reflective refinement and representation of data: a timeline of evaluation practice

A reflective turn of data analysis was critical to advance from preliminary findings, emerging from the main themes, to a final refinement of findings. This was achieved by a combination of memos and learning journal writing with a visual representation of findings in the YPET. Through the tool of memo writing in ATLAS ti, I wrote freely and extensively my interpretation and reflections about the possible meaning of a set of connections, patterns and preliminary findings. Figure 5.8 shows the list of all memos produced.

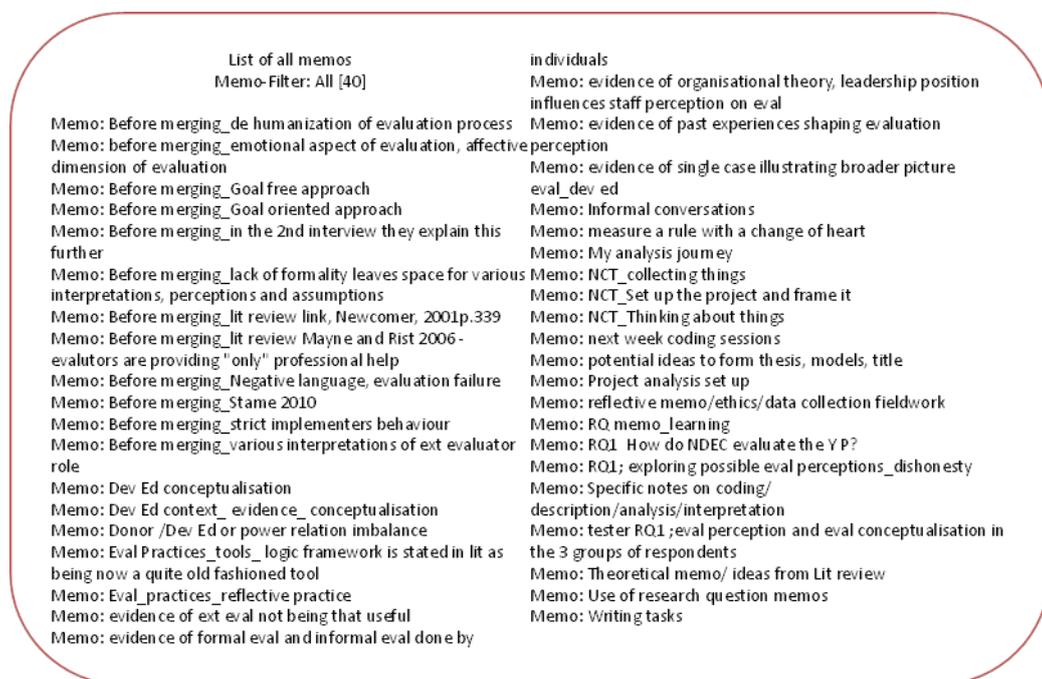


Figure 5.8:List of memos written during the analysis in ATLAS ti.

From here, I iterated with my learning journal, in which I had jotted field notes, helpful to illuminate further refinements. For instance, with regard to the connection between space and practice, I could not see any clear connection through the query tool in ATLAS ti., but in the journal I had written an extensive entry about the impact of NDEC's change of office to a smaller space.

This realisation ensured that ATLAS ti. has massively supported the organisation and coding of my dataset, but a final refinement of analysis was computer-free, but brain-assisted, in which I reflected on how findings could be represented over time in the YPET. This was another stage of interpretation, through the descriptive portrayal of the main themes emerged, in which I drew mind maps of the first preliminary account with five themes emerging as main findings: power relationships; perception of evaluation; participation in evaluation; reflection, and co-occurrence of practices (see Appendix 7.b; refinement of findings using brain-assisted methods). As results of reflection and intensive writing retreat sessions (explained in section 9.6), these five findings were refined into three: participation in evaluation, power relationships, and co-occurrence of EP with other working activities.

In analysing how these final findings intersect, a foundational timeline of YPEP was generated – youth project evaluation timeline (YPET) - it was adapted from previous DSP studies to explore instances of change over time (Shove *et al.*, 2009, 2012) - and it represents my interpretation of the patterns of change and readjustment of EP; the three findings give fresh insights for gaining an in-depth understanding of the circumstances of change in EP in a DE setting (see Figures 6.2, 7.1, and 8.1). Another initial version of the YPET was generated to chronologically situate the factual evidence of the dataset – the youth project chronological timeline (YPCT), as visualised in Figure 6.1. A final version of the YPET combines the previous timelines to illustrate an overview of how evaluation practice changed across the YPET (see Figure 9.3).

The use of the YPET allowed for the exploration of change in EP, in terms of the altering relationships between their constitutive elements: material, competence and meaning. The first was depicted by observing practitioners' attitudes towards M&E tools, such as the project outcomes matrix, logic framework and external evaluation report; the second, by observing the skills that allowed practitioners to participate in the evaluation; and the third, by the relevance practitioners assigned to evaluation regarding their aspirations, expectations, and visions. The portrayal of the EP through the YPET emphasised the particular arrangements of these elements and how EP may evolve as they change. These changes might have been influenced by a specific event or PCE, as illustrated in Table 6.1.

The YPET, as a rich longitudinal tool, captured objective data, such as the date of project approval and milestones for evaluation deliveries, as well as subjective data such as, the incidents with practitioners or change in their relationship with others. As the analysis came to an end, various attempts were made to design the final timeline. Although it was challenging to affirm that the research analysis was finalised, there was the need to be a cut-off point; so for the purpose of answering my research questions, this was a 'provisional stable' point (Saunders *et al.*, 2005), with room for improvement. This final stage of analysis was vital to place the findings in a longitudinal perspective to draw robust answers to my research questions.

A final note is about the in-depth richness of the single case that was made clear with the emergence of two unexpected findings: the stage of project design, as deeply influential to how evaluation was subsequently practised (see section

6.4); and the role of a retrospective reflection, facilitated by the research process, in increasing the awareness for the need of a practice-based evaluation (see section 8.3.3). The robustness of my research design allowed me to see beyond what was initially planned, for example, in responding to ‘how’ stakeholders evaluate the project, alongside the initially ‘what’. This unexpected emergence of rich data supports the need for more in-depth qualitative cases, exploring the particularities (Simons, 2015) of a phenomenon that may have been already looked from a broader perspective, as it was EP.

In summary, the journey of analysis started with an iterative approach (Huberman and Miles, 2000) focused on descriptive and process coding (Saldana, 2009), through the use of ATLAS ti., (Friese, 2012), to allow the transformation of data (Wolcott, 1994; Simons, 2009) – (see Table 5.13; influential approaches to my analysis and interpretation). It then moved on to a progressive focus of analysis through the query of the dataset to draw preliminary findings from a longitudinal analysis of how evaluation had been practised and changed over time (Saldaña, 2002). Having used a timeline as an interpretative tool to portray the trajectory of EP over time (Simons, 2009), it was possible to refine five preliminary findings into three; as well as to gain fresh insights to unpack change in EP (Shove *et al.*, 2012). A final step was to design three timelines, one per finding, to explore how each evolved over time; then compare them with the main timeline to elaborate answers to this study’s research questions. The summary of my dataset in the form of three timelines of EP supported the process of understanding emerging data through the theory of DSP (Shove *et al.*, 2012).

5.3.8 Ensuring data quality

Interpretative qualitative scholars have been working to ensure that data quality can be valued, not only through scientific strategies, but by applying alternative approaches (Wolcott, 1994). Guba and Lincoln (1988) propose a parallel criteria to the conventional internal and external validity, reliability, and objectivity that include trustworthiness and authenticity, which has been used in case study research (Simons, 2009). Trustworthiness is the parallel term to reliability to justify the value of an account as more likely or more credible (Wolcott, 1994). Authenticity refers to the “*fairness*” (ibid.) through which research participants were treated and their accounts translated. These terms connect with my social constructivist approach to qualitative enquiry, because they are concerned with how data is “*negotiated and understood*” (Simons, 2009, p. 128), and have been widely used in empirical studies on EP. For this reason, I followed Simons’ (ibid.) procedures of trustworthiness and authenticity to ensure the quality of this research, which is explained through data triangulation and response validation. This section also includes a note on generalisability from a single case.

5.3.8.1 Data Triangulation

Triangulation is defined as “*a means of cross-checking the relevance and significance of issues from different angles to generate and strengthen evidence in support of key claims*” (Simons, 2009, p. 129). Although some scholars have not agreed with the literal use of the term to ensure the quality of qualitative research, arguing that it puts data into confinements (Guba and Lincoln, 1988), others have adopted a balanced position about its meaning and used authenticity

to claim the quality of their studies (Simons, 2009). Denzin and Lincoln (2005) pointed out four types of triangulation: methodological, investigator, theory and data triangulation. The latter is adopted to defend the quality of this study, because the preoccupation to gather various perspectives about the phenomenon of evaluation was visible from the beginning of this research. For this reason, I interviewed three groups of participants, whilst also having informal conversations with other informants, such as volunteers and other practitioners. This triangulation reinforced that the main views emerging from data collection were authentic and not potentially biased by my presence in the field. Also, data triangulation was ensured by using various sources to explore the phenomenon under investigation, combining interviews with observation and documentary evidence, which allowed me to monitor the trustworthiness of this case study throughout.

5.3.8.2 Respondent validation

Respondent validation means to “*check the accuracy, adequacy, and fairness of observation, representations and interpretations of experience with those whom they concern*” (Simons, 2009, p.131). In turning to my research, three notes appear relevant:

1. Two phases of data collection

As illustrated in Figure 5.6 I collected data from three groups of participants during the implementation and the reporting of the evaluation, which allowed me to interview four participants twice. This method not only refined the preponderance of some issues at the expense of others, but also made possible to validate previous ideas shared during the first interview. Moreover, the

interview protocol that followed in the second interview was based on the previous transcript from the participants' first interview, which for some interviews needed clarification for accuracy or additional rephrasing (see Appendices 4.b and 4.c; first and second interview protocols). This was extremely useful and in most cases I checked and clarified that the key messages taken were accurate, but in one case, one participant actually clarified one specific association, that I had unintentionally made.

2. Informal conversations

To ensure that the main issues emerging from data collection were representative of these groups' experiences, I held various informal conversations along the way, which were very spontaneous and most times started with: *'How are things going? Where are you up to with the evaluation?'* Frequently, participants were the ones who reinforced key points discussed earlier in our interviews, without my direction to a specific issue. Two examples in different stages of evaluation were with Tyler, practitioner, who shared their thoughts about why evaluation was being difficult for their colleagues. In this conversation, Tyler repeated the same reason mentioned by them, so I was reassured that I had well understood their message. The second example happened during the reporting stage of evaluation when another practitioner, Jesse, confessed to be job-seeking, because they could not handle the lack of strategic direction of the organisation. In the previous interview colleagues had shared their concern about the overall negative perception of evaluation and how that was affecting their motivation as project workers.

3. Sharing preliminary findings with research participants

Although some qualitative scholars suggest that respondent validation is best made through the sharing of the interview transcripts for participants' validation (Huberman and Miles, 2000; Silverman, 2011), this study followed a slightly different route. Due to the longitudinal dimension of the study, I started transcribing the first phase of interviews immediately, to ensure that I would have most of it in place to book the second phase of interviews. As explained earlier in this section, the informal conversations were reassuring regarding the trustworthiness of my data collection, so I kept my transcriptions flowing until an advanced stage. For this reason, I decided to use an alternative approach and to share a short summary of preliminary findings, inviting participants to add their views, either to agree, disagree, or comment (see Appendix 7.c; summary of preliminary findings shared with participants). Another reason that reinforced the alternative of sending this summary, instead of the original transcripts, was the constraint of a part-time and self-funded mode of study, so I had to account for that when making these and other research decisions.

I sent the summary to twelve participants who took part in my research and I received four long and well-structured comments and feedback. Surprisingly, not only did the summary of findings resonate with them, but also they added additional preoccupations on which they had been reflecting and would like to see addressed. The rapport with my participants was developed during a long period of field work and their reflective practice appeared to have been prompted by the research process, which ensures the fairness of how they were treated and the authenticity generated in the process of data collection (Simons, 2009). Possible explanations regarding the low number of responses, possibly,

confirm one of this study's preliminary findings - the co-occurrence of practices in a DE context - as practitioners rarely have spare time to participate in *ad hoc* activities such as the further involvement in this research. Alternatively, although the low number of responses can be seen as a potential limitation of ensuring the quality of the study, it reassured me that I had taken the right decision, because participants would have been less likely to respond to the lengthier transcript checking either.

5.3.8.3 Generalisability in case study: interpreting in context

Although open to debate I argue that this single case study can be generalisable from its "*particularity and uniqueness*" (Simons, 2009, p.164), however it is relevant to elucidate that this thesis does not aim to generalise to the theory, but rather ensure that the inferences I draw from my data may be applicable in other particular contexts. As such, and bearing in mind the context of this study and the aim of exploring how evaluation has been practised in a single DE setting, my responsibility is not necessarily to generalise (as the study has not been designed to do so), but to demonstrate how, and in what ways, my findings can be applicable to other contexts and used by others. As Simons (2009) advised, generalisation in case study research is not dependent upon a propositional statement that either applies to a wider context or not, thus case studies are not typical in the way that the inferences made can vary. Case studies are particular and unique and, in this regard, the transferability to other settings is the "*reader's burden*" (Mertens, 2005, p.271), as they may generalise on the basis of recognising similarities and differences to their own experience.

A possible limitation of the single case, as a research design, regards the possibility of generalisability, in the sense of judging the quality of the study by the transferability of the findings to the wider context. As this is not the understanding of generalisability that I hold, this limitation was addressed by following some qualitative scholars' procedures applied in previous evaluation studies such as, particularisation and concept generalisability (Guba and Lincoln, 1988; Stake, 2005; Simons, 2009, 2015). In this thesis, the issue of generalisation has followed two protocols: concept generalisation and in-depth particularisation that are detailed in chapter nine, because it supports the claim that conclusions can be drawn from these research findings.

5.3.9 Ethical considerations in the fieldwork

There is broad literature related to the need for ethical considerations in social science research (Laine, 2000; Guillemin and Gillam, 2004; Denzin and Lincoln, 2005; Shaw *et al.*, 2006; Silverman, 2011). In terms of the meaning of ethical considerations, Silverman (2011) consider ethics the act of posing the theoretical question of how one ought to live, and morality an answer to that question – a way of living that is regulated by rules, rights, duties and commands. From a researcher's perspective, ethics signals the necessity of conducting research without any philosophical or ontological biases (Simons, 2009), and in this respect two issues are addressed: first, confidentiality, anonymity and use of information; second, my role as a researcher. The confidentiality of the research, after having followed the ethical guidelines set out by Lancaster University and received their approval to proceed with the fieldwork, was officially ensured by the participants' completion of the

informed consent form. Alongside this informed consent, I had previously sent a research invitation letter to all participants, summarising the aim of the research and their role in it (see Appendix 4.a; research invitation letter). Regarding anonymity, a gender-neutral pseudonym was used for research participants, as well as a general description to anonymise the DECs organisations. From the ethics research committee's point of view, this research was considered low risk, as it did not include contact with vulnerable people. It was also essential to ensure that the research topic was not sensitive and there was no likelihood to cause harm or distress to participants. Secondly, my role as a researcher had to be acknowledged, particularly, in relation to my previous volunteer experience in NDEC. I particularly remember having to politely refuse one participant's request to help them in interviewing some beneficiaries. That was an embarrassing experience for me as a foreigner and a non-native English speaker, because I felt the need to give them an extra explanation to ensure the access to data collection, but I was unsure about the ethical code to do so. I opted for being honest and took advantage of the opportunity to clarify my role as a researcher, reassuring that if I had not been researching their organisation, I would have been available to help. This example illustrates one ethical issue that I had to re-negotiate throughout the research journey, as well as the construction of a trustful atmosphere between me, as a researcher, and my participants.

5.4 Reflecting on the research journey

5.4.1 Methodological, technical and other issues encountered

From a variety of issues encountered, I describe five methodological and three technical issues, as visualised in Table 5.16 (overleaf):

Methodological issues	Technical issues
1. Research foci – from evaluation process to progressive focus on practice	1. Sense of participants’ manipulation during the research interview
2. Decision about how to position the research in the field of evaluation rather than in the DE domain.	2. Evaluation form: funders excluded from filling in the form and possible ambiguity regarding the period covered by the form.
	3. Unequal depth of analysis of the data between the YP partners.
3. Difference about conducting and researching evaluation	4. Different role of writing and speaking throughout the research
4. A SPT lens focus on individuals’ practice and I had to analyse participants’ evaluation practice from a group’s perspective instead - to ensure the confidentiality and anonymity of the participants.	
5. Encounter a DSP analytical framework at an advanced stage of the research.	

Table 5.16: Methodological and technical issues encountered

5.4.1.1 Methodological issues

The first methodological issue faced concerns the foci of the research. For example, while the initial literature reviewed appears to indicate an exploration focused on the construction of the evaluation process, the advance of the literature search combined with a pilot study confirmed the need to focus on ‘practice’ as the main case, rather than ‘process’. This also meant a redefinition of the literature review and the general approach to data collection and analysis. A second methodological issue was the decision of positioning this thesis, primarily in the field of evaluation, rather than in the field of DE. This decision became clear during the analysis, where two bodies of data were emerging as distinct, one was the EP, as the phenomenon under investigation and the other the domain of DE, as a research setting, serving as an interesting context to

observe EP over time (Patton, 2002). As a PhD thesis has specific limitations, I knew that I had to narrow my focus once more and, in reflecting about the initial motivation to conduct research, I knew that I wanted to position this study within the EP domain (again, a matter of ‘researcher loyalty’, described in section 5.2.4). This decision made sense, as I have been connected to evaluation as a practitioner and would like to pursue my career in this field.

A third methodological issue was about the difference of conducting and researching evaluation. This issue arose during the data collection when I felt the constant need to reassure participants that I was neither judging their competence nor their organisation’s performance, as the aim of my study was to research how evaluation has been practised within a DE domain. So I was researching EP and not conducting an evaluation of the YP. To overcome this challenge, before starting each interview, I read an introduction of the research to the participants, reminding them about this difference.

A fourth methodological issue was the apparent limitation that arose from the use of a SPT lens that focuses on an individual, rather than in a collective practice. Due to ethical reasons I analysed EP from a group’s perspective (embedded cases) for two reasons: the first concerns the case study design regarding the presentation of the composite profiles of participants (described in section 5.3.5), which allowed the reader to make connections with their own setting (Simons, 2009). Moreover, I believe this intention was achievable from a recollection of accounts, in which I drew attention to some aspects rather than others. However, the reader holds the freedom to pinpoint a particular nuance, without disclosing the individuals’ anonymity. As such, the thesis is interested

in the collective practice of evaluation within an organisational setting, which was accomplished through a SPT view of evaluation, from each group, and as a whole. The second reason to justify my decision for analysing practices collectively, relates to the connection between a DSP analytical lens with the literature on communities of practice (CoP). In this respect, previous research on CoP suggests that the way individuals interact through practices needs to be explored in the context of narratives, both personal and collective (Wenger, 1998). So even though DSP has been mainly used to explore individuals' practices (Shove *et al.*, 2012), my research considers how individuals practise evaluation collectively and how that practice is negotiated and perceived across different groups. Besides, the richness of this data set was the combination of personal stories collected over time, to produce an in-depth collective account of how participants practised evaluation and how it changes over time. This account was organised according to each participant's professional affiliation, because literature on EP (Saunders *et al.*, 2011) has recognised the need to advance the current knowledge about each professional groups' challenges in evaluation.

Lastly, a fifth methodological issue arose with the selection of the analytical framework employed. As I subscribed to the view of evaluation as a social practice, the application of a SPT framework was decided from an early stage; however, the encounter with the specific strand of the DSP framework (Shove *et al.*, 2012) occurred later in the research journey, specifically, at the end of data collection. Before this encounter, I had considered the use of RUFDATA (Saunders, 2000) as a potential framework, but because it represents a reflexive tool to plan evaluation, there was the risk of deductively imposing it on the data,

ending up checking how well NDEC was using it or not. In other words, my research aim was to describe what practitioners did when evaluating the YP - and how that practice changed over time, so for this reason a DSP framework seemed appropriate to conduct this research study's analysis.

5.4.1.2 Technical issues

Various technical issues occurred throughout the research journey, but for the purpose of the thesis I outline four. The first issue encountered was some participants' manipulation of the research interview. As they were asked to identify relevant episodes in the YP evaluation, one participant had prepared in advance written hand-outs for the first interview. This detail felt uncomfortable, as noted in the following extract from my learning journal:

During most of the interview the interviewee picked up mainly their ideas, already written in the handout given to me, at the beginning. So, I constantly had to prompt them, sometimes successfully, other times not so much. They seemed very protective about what they had written so in a sense very well-aligned with the hard-evidence paranoia that they seemed keen to defend. In the second interview, although papers have not been given, they did refer back to those, by explicitly saying, "*I think I already said that Joana. I even wrote it down on the papers given to you*".

December 2012

Extract 5.1: Extract from my learning journal, December 2012

This situation was manifested as manipulation from a research participant, but also produced further reflection about my bias, especially after the transcription of other interviews, where participants mentioned additional issues regarding this specific participant, as the following example illuminates: "*They came in and, sometimes, I think they would come with their own agenda*" (Sam). As a researcher, after reading Sam's transcript, I noticed that I felt the same about this

‘manipulative’ participant – who seemed to have their own agenda. From that moment onwards, I strictly monitored my bias, particularly in relation to this participant (Simons, 2009). This made me develop a profound reflection in the remaining journey.

A second technical issue occurred during the first phase of interviews when I asked coordinators and practitioners to fill in a form with various evaluation activities. Funders were excluded because I assumed that these activities were not part of their role (see section 5.3.7; composite profile of the research participants). On the one hand, data on funders’ activities was collected from other documentary sources of evidence alongside their interview transcripts; on the other, it may well hide some distortion, and for that reason I approached it as a technical issue.

A third technical issue encountered was the realisation that I have not been able to analyse data from NDEC’s partner in the same depth as I did with them, because, whilst the longitudinal dimension was secured in the NDEC case, their partner’s participants were not available for a second interview. Thus, some of the data related to the partner was collected through other peoples’ narratives (Weiss, 1995). Ethically speaking, I decided not to analyse that part of data in-depth, despite its coding and categorisation, because some turned out to be potentially sensitive regarding the confidentiality of participants. For the purpose of the thesis, part of this data was only used as background information (Patton, 2002), and sensitive evidence is in a confidential appendix (only made available for examination purposes).

A fourth and final technical issue was the realisation of the different role of writing and speaking about my research. An illustration was when I presented my research in a postgraduate conference held in Lancaster University, and the audience asked me at various times to clarify the meaning of DE as a sector, and not even once did they mention EP. In attempting to bring the audience back to my research focus - EP - I realised that the way I was speaking about my research was considerably different from the way I had been writing about it. Peer feedback and support was vital to align speaking and writing, and was made through the adaptation of my regular writing retreats, by incorporating a half hour 'talking space' to develop presentational skills (see section 9.7; writing retreat as a daily practice). Having summarised the main issues encountered, the reflective space next describes two stories from the fieldwork.

5.4.1.3 Reflective space

During one interview the telephone in the office rang and I focused my attention on the interviewee, assuming that they wouldn't pick it up but, surprisingly, they ran upstairs, saying that there was nobody else in the office. On their return, I managed to get the interview's pace back, when the telephone rang again. Luckily, this time the interviewee said, "I think I will leave this one", and I ironically thought to myself: 'that is very kind of you'. After this situation, I reflected on the expectations researchers have, which requires that the world around us should stop to support our research process. This episode made me reflect about the different meaning attached to the interview, by this participant and myself. For me, it was absolutely fundamental that the interview was not interrupted; I had even put a note on the door advising that audio-recording was in progress. For this participant, the real life kept going and the telephone had to be picked, the office has to keep running, and if they managed to be interviewed at the same time, splendid.

October 2012

Extract 5.2: Extract from my learning journal, October 2012

Today in the NDEC's office I noticed my name written on a wallboard to be included in the team supporting one of DfID's reports. Rather puzzled, I felt embarrassed to clarify my current role as researcher, and the need to alter the previous agreement as a volunteer supporting their EP. That had to be made explicit, again and again. One strategy that I found useful was to send practitioners some literature on EP every time they forget about my researcher's role; which was a lot. The need to remind them about the new boundaries of our relationship was frequent during the first phase of data collection, which prompted my discomfort, when I spotted my name as one of their colleagues, rather than an independent researcher. These moments of reflection were confusing, because on the one hand, I wanted to be polite and ensure they trusted me, in order to conduct my fieldwork until the end; on the other hand, I realised that participants were counting on me to do something ethically doubtful. Also, in the research invitation letter, I had formally advised them about the restriction of helping with evaluation, even offering help on other relevant issues; however because I had been a practitioner in the past, I knew that it was hard to keep the balance of respecting another's role when there were no other *pro-bono* alternatives. This is so much the case, when the funder's pressure was on and the clock kept ticking. As I faced this dilemma of making my researcher's role explicit, I tried to avoid giving unnecessary recommendations, ensuring participant's clarification about my role.

January 2013

Extract 5.3: Extract from my learning journal, January 2013

5.5 Summary and conclusions to the chapter

This chapter introduced, described and discussed the methodological decisions supporting the thesis. Research questions will be answered by employing a single in-depth case study exploration. The focus is on exploring how evaluation is practised in a DE setting, how it changes over time and the influences to that occurrence of change. This reflects the intention of gaining an in-depth understanding about the meaning assigned to evaluation by research participants, which will be ensured by a longitudinal data collection based on multiple data sources – interviews, observation, documentary analysis, and informal conversations. A descriptive account of the research methods used was discussed, alongside the procedures adopted to analyse and interpret the dataset.

Lastly, a reflective account of the methodological and technical issues was presented. This summary is illustrated in Table 5.17 (overleaf).

The following part two progresses this thesis by presenting a descriptive account of the findings and its respective discussion across chapters six, seven, and eight. The decision was taken to aggregate data and analysis for each finding, to enhance the flow of the account.

Approach	Single in-depth case study to explore the EP particularities of a three-year project and how it changed over time.	Section 5.2
Methodology	A SPT framework was employed for data collection and analysis. Longitudinal data collection was based on multiple data sources – interviews, observation, documentary analysis, and informal conversations. In total, 16 interviews; 118 hrs. of observation (94 hrs. of participant and 24 hrs. of non-participant observation); 30 documents and 8 informal conversations were conducted in two phases of data collection. For data analysis, a framework of DSP was used to analyse the trajectory of EP over time. Interpreted data was displayed in a series of timelines, according to each finding.	Section 5.3.2 and 5.3.6
Research setting and sampling	Purposive sampling was used to find adequate organisations to fit the research criterion. In doing so, one of the six DE organisations initially surveyed stood out as an information-rich case (Patton, 2002). The research setting is the YP evaluation and 12 individuals were invited to be part of the current research on the basis of their involvement with the evaluation.	Sections 5.3.3 and 5.3.4
Data analysis and interpretation	Computer-assisted thematic analysis based on the full transcription and initial immersion into the data, descriptive and process coding, the query of data in ATLAS.ti and the reflective refinement of findings, in which a timeline of EP was designed to discuss and display the intersection of findings.	Section 5.3.7
Limitations	A small sampling may have limited the inclusion and selection of more case studies, which would have allowed further comparisons. A retrospective data collection may have obscured the interpretation of data, as past events were shared by participants from their current standing point. The decision to adopt a DSP analytical lens that had not yet been applied to the domain of evaluation might have been risky, in the sense that the depth of exploration had no previous criteria to check against.	Section 5.4.1

Table 5.17: Summary of methodological decisions

6 PARTICIPATION IN EVALUATION PRACTICE

6.1 Introduction and structure of the following chapters

This part of the thesis deals with the main research objective: to explore how development education (DE) organisations evaluate their interventions within a funder-recipient based context. The following chapters (six -eight) provide both a descriptive and analytical account of the case's evaluation practice through the discussion of the main issues that emerged across the trajectory of the youth project (YP) evaluation. By analysing these issues in light of the literature reviewed (chapters two-four), this part combines the presentation of findings and related discussion. These chapters contain references to three elements, as follows:

1) A summary of the key chronological stages of the project EP, represented by a chronological timeline based on factual documentary evidence obtained from the organisation's archive, interview transcripts and field notes from my observation and learning journal (see Figure 6.1).

2) A list of practice-change episodes (PCEs), which are based on participants' accounts shared during the interviews and field notes from my observations and learning journal. Informed by Houghton (2003), I use this term to describe an instance of my account that emerged as influential to the change of the YP evaluation practice (EP). The PCEs presented in chapters six to eight are visualised in Table 6.1 (overleaf):

Thesis chapter	Name of the PCE	Brief description	Stage of evaluation
Six	1. Practitioners in the design of the YP	Practitioners did not participate in the design of the YP	Design of the project
Six; seven; eight	2. DfID's change of goals	Review of DfID's evaluation mandate and subsequent change of their goals	Planning Evaluation
Six	3. GMA's appearance	GMA is appointed on behalf of DfID to monitor and evaluate projects funded under the DAF scheme	Planning Evaluation
Six; seven; eight	4. Funders' different evaluation approaches	Participants realised contrasting evaluation approaches between DfID and GMA	Planning Evaluation
Six; eight	5. Absent coordinator	Absence of the NDEC's coordinator on leave	Planning Evaluation
Seven	6. Film evaluation	Practitioners asked permission to funders to produce a more informal evaluation report through a film evaluation; which was refused due to the performance-evidence approaches taken.	Reporting evaluation
Seven	7. External evaluator (EE) arrival	EE comes in to support the process of planning, implementing and reporting evaluation	Three stages of evaluation
Eight	8. NDEC's extraordinary strategic meeting	An extraordinary meeting was held to decide about the future of the organisation.	Reporting evaluation
Eight	9. NDEC's change of office	Due to the Government's austerity measures, NDEC had to re-organise their financial provision, which led to a change of office to a shared open space.	Reporting evaluation

Table 6.1: Practice-change episodes discussed

3) A timeline representing, if and, how each finding changed across the EP over time. These are based in the application of a dynamics of social practice (DSP) lens to analyse the project trajectory of EP (See Figures 6.2; 7.1, and 8.1). In doing so, I differentiated the background evidence from my interpretation of data.

Chapter six to eight represents the main issues that emerged from findings during the data analysis – participation (PAR); power relationships (PR); and co-occurrence of evaluation practice (COoEP) with working activities (WAs). The structure of this part was based on methodological and practical decisions. First, a methodological decision of alignment with the analytical framework of DSP (discussed in chapter three) was to separately explore each issue that influences how EP change over time. The presentation of findings separately supports the subsequent discussion of the how they intersect in, and with, EP over time. Also, the separate exploration of these findings enables an in-depth understanding of their influence in the practice of evaluation. Second, a practical decision was the balance of having three shorter chapters rather than a lengthier one. Following Shove *et al.*'s (2012) procedures, the decision was to explore evidence and discuss each finding separately to enhance how the discussion answers my research questions. That resulted in three chapters representing issues around participation, power relationships and co-occurrence of EP with WAs, respectively. However, this decision has presented the challenge of the sequence of the 'story', so in this respect, I followed a chronological sequence, and started the descriptive in-depth account at the point of my entrance to the fieldwork. I then move back and forward to retrospectively tell the reader important background information and past events experienced by participants, because I wanted to gain in-depth insight of the changes occurring. For this reason, I described each issue individually across all stages, because it is these dynamics that allows change in EP to be captured over time.

In terms of structure, the three chapters each deal with one separate issue. I acknowledge that I could have organised the structure of this part by the stages of evaluation, but because the timeline already visualised those, I decided to organise it by the relevance of findings that emerged. Each chapter starts with a brief introduction and a timeline locating the occurrence of that issue within the YP chronological trajectory. This is followed by a descriptive account of findings through which the issue emerged. The discussion of findings interweaves with the descriptive account, and this strategy was useful to avoid repetition. For instance, in this chapter six, the timeline is labelled ‘Participation in the youth project evaluation timeline’ (PAR-YPET) and the intention is to visualise how the issue of participation influenced the change in the project EP. The acronym is replicated for the subsequent chapters, with PR-YPET and COoEP with WAs-YPET, for power relationships and co-occurrence of EP with WAs, respectively. The display of findings through a timeline is an interpretative technique previously used to explore change in social practice (Shove *et al.*, 2012). Table 6.2 (overleaf) illustrates the research findings and issues emerged.

Thesis chapter	Issues emerged	Research findings
Six	Participation (PAR)	1) The circumstances under which a DE project is designed influences the participation of stakeholders in the evaluation of that project. As a result, the issue of participation influences how evaluation is planned, implemented and reported – practised.
Seven	Power relationships (PR)	2) Power relationships are influential to the way evaluation is perceived and practised by DE practitioners; different power relations shape evaluation differently.
Eight	Co-occurrence of EP (COoEP) with other working activities (WAs)	3) Evaluation is a collaborative and collective practice that co-occurs with other working activities. The competence of the DE coordinators and practitioners to evaluate is influenced by a COoEP setting, as well as influences how EP changes over time, particularly during periods of unpredictable circumstances (DE working environment as a context where EP concurs with other working activities).

Table 6.2: Research findings and issues emerged

This overview ends with Figure 6.1 (overleaf), indicating the chronological facts across the youth project evaluation timeline (YPET).

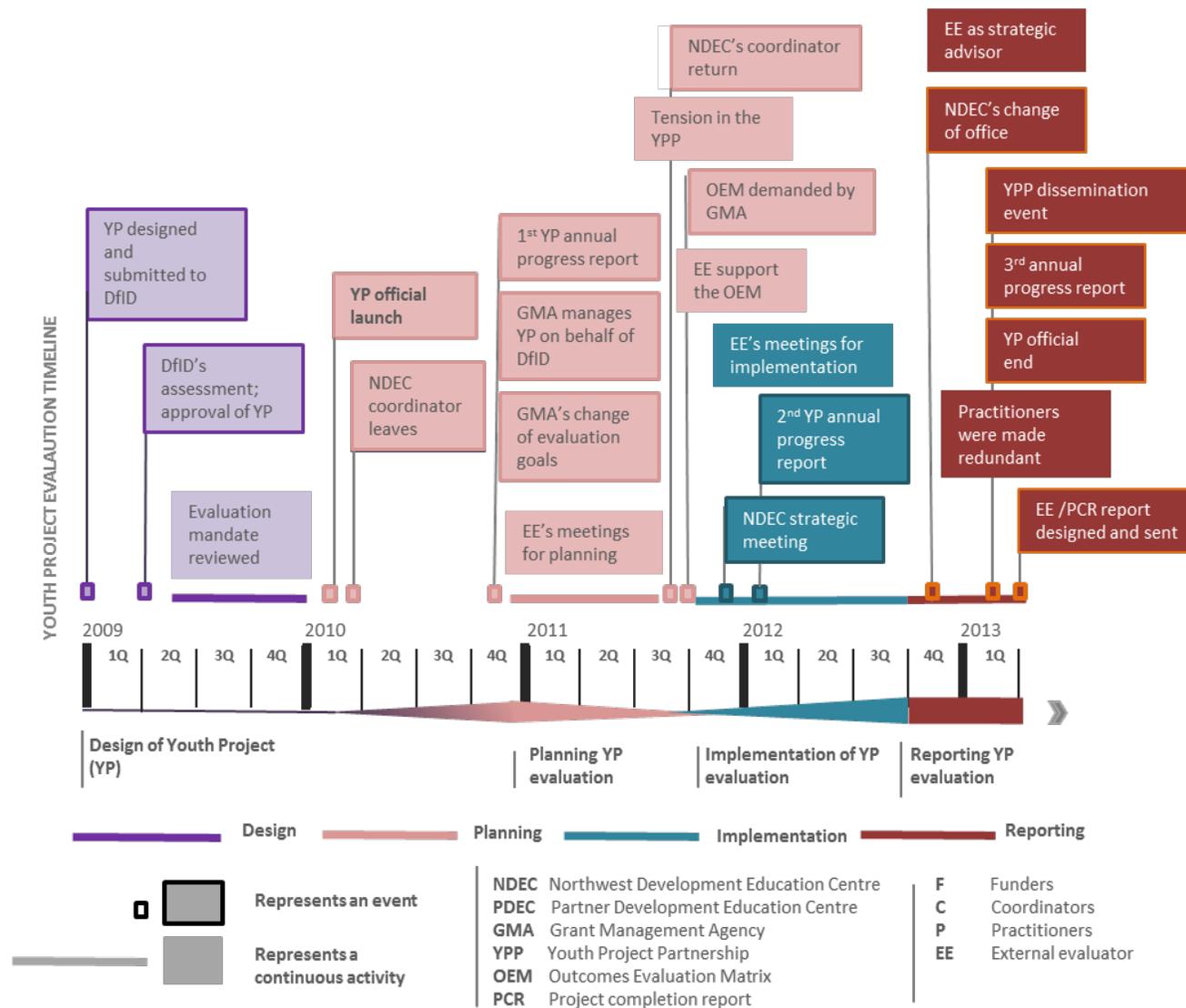


Figure 6.1: Youth project evaluation timeline: chronological facts

Structure of this chapter

Participation was one of the three main themes emerging from the analysis and because it was identified by participants as having had a key role during the initial design of the project, it is the first theme presented and discussed (see Table 6.2; research findings and issues emerged). The chapter starts with an explanation of the rationale for structure the next three chapters (section 6.1). Section 6.2 contextualises the YP background, examining the role of a funding-dependent context and its implications for the competence of practitioners. Section 6.3 examines the barriers to participation depicted by the participants and how these influence their practice of evaluation. Section 6.4 outlines a typology of practitioners' engagement in EP over time (disengagement, forcing, resistance, coping, and collaboration). Section 6.5 describes the change observed in the YPET, explaining what I mean by the paradox of participation. Section 6.6 summarises the findings and conclusions of the chapter.

Next, this chapter presents a visualisation of the issue of participation and its emergence across the evaluation timeline (PAR-YPET). The rest of the chapter unpacks Figure 6.2 (overleaf). For clarity purposes, 'the YP evaluation' is referred henceforth as 'the evaluation', and the same applies to 'the YP project', referred as 'the project' and 'the GMA' is referred to as 'GMA' (for the purposes of anonymity, I altered the name of this funder to Grant Management Agency (GMA), so it becomes a proper noun).

FINDING: Participation (PAR)

- BPP** Barriers to practitioners' participation in YP design
- LTPE-EP** Typology of practitioners' engagement in evaluation practice
- C-EP** Change observed in evaluation practice: by adaptation
- PP** Paradox of participation

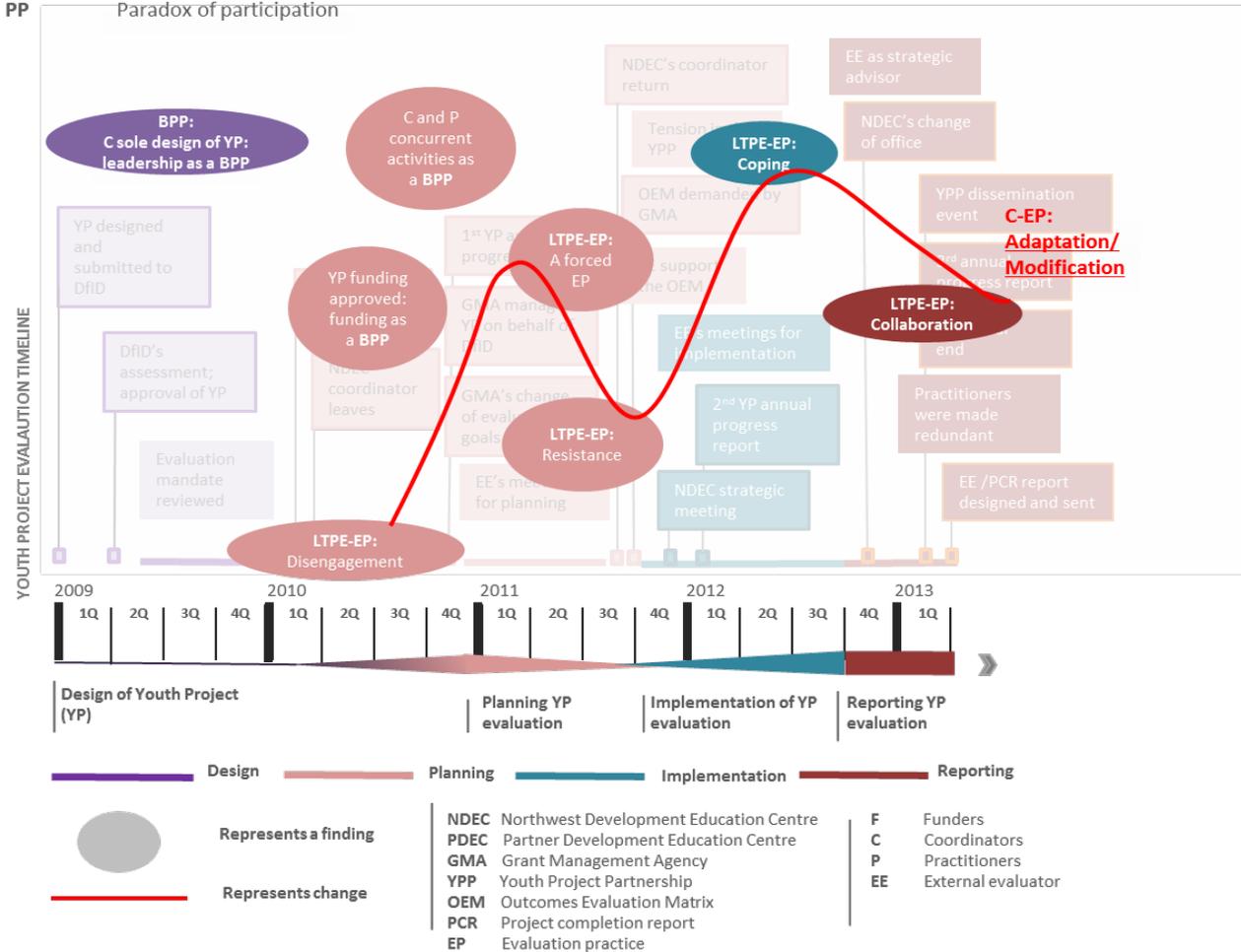


Figure 6.2: Participation in the youth project evaluation timeline (PAR –YPET)

6.2 The context of the youth project: background information

The Northwest Development Education Centre (NDEC) has been involved in various types of learning experiences with the intention to design a DE project to be funded by the Department for International Development (DfID). As the youth project (YP) appeared to fit in one of the funders' schemes - Development Awareness Funding (DAF) - a proposal for a three-year intervention, in partnership with another DE centre, was submitted in the beginning of 2009, and officially approved in the last quarter of 2009. The project was then launched in early 2010.

At that time, the Coalition Government's decision to review the DE policy on funding coincided with the starting of the project implementation, alongside the planning of its evaluation. That political review initiated a cycle of change for the sequence of DE project management, particularly affecting the project EP. In the midst of internal and external constraints recalled throughout this account, a critical episode was DfID's change of their parameters without further notice, through the announcement that funding could be withdrawn if some DE projects were unable to demonstrate evidence of outcomes achieved (see Table 6.1; practice-change episodes discussed). Participants shared that this change seemed to have severely impacted other DE organisations, in particular the NDEC. For instance, the unexpected withdrawal of funding occurring with other DE colleagues seemed to have had a highly disruptive effect for the team, because that abrupt change meant that even the projects already funded, as it was the YP, could see their money withdrawn. Consequently, various perceptions of evaluation were shared, as some misunderstanding about the need for evaluation was perceived amongst the team, as Jesse, practitioner, described:

I think that one of the key things about evaluation is that we all understand the need for it in an academic, business, and educational level [*pause*] **but the way we feel about it is very different, until our fear changes.** (Jesse, 2nd interview, **emphasis added**)

The context of DfID's abrupt change of parameters have increased participants' notion of fear, especially because they were unexpected. For instance, the NDEC coordinator had developed a relationship with DfID, before the project submission, so was not expecting all these changes and even an eventual closure of some other projects. Before these political changes the organisation used to receive, implement and evaluate mini-grants under a flexible approach, as Sam, practitioner, explained:

In my experience, DFID's funding had always been a bit more flexible and you would write your own evaluations and **whatever you would tell them, they would take it as given** and we did not have to provide as much evidence, as they wanted this time. (Sam, 1st interview, **emphasis added**)

DfID's previous approach to evaluation was activity-driven rather than outcome-driven; which meant that practitioners could present a brief evaluation on each activity of the project, unconcerned with the demonstration of how these activities have contributed to achieve the planned outcomes. In this context, practitioners were used to evaluate under this activity approach; however, as Sam emphasised, one day funders unexpectedly changed their mind and a set of new practices arose, alongside the threat of losing the funding. The consequences of the DfID's change of approach were felt in the confidence of the YP team, as Stevie, practitioner shared:

From the big things that we seem to have as barriers [*to understand and practise evaluation*] one is lack of knowledge and consequent to that, is a lack of confidence to have the knowledge.

Another barrier is that we do not have access to affordable continuing training. (Stevie)

This was early 2011 and the team had to plan the evaluation, but the project design stage seemed far too distant to remember their thoughts about project goals and outcomes achievement (see section 6.4.1; disengagement). In funders' terms the stage of design has been referred to as 'formulation' of the project, which normally includes designing a technical and financial proposal; accounting for management arrangements; having a monitoring and evaluation (M&E) plan, and receiving a financing decision (EuropeAID, 2004). For the purpose of the thesis, I refer to 'project design', because it is the term most frequently used in the literature on EP. Next, I present participants' articulations of the project's context as funding-dependent, versus their competence in practising evaluation.

6.2.1 Funding dependency versus competence

Overall, the changes made by the Coalition Government seemed to provoke fear and uncertainty within this DE context and that was the scene when the project started. In terms of management, I invited participants to describe NDEC's strategy for financial sustainability, the measures that were in place, and the alternative sources of funding for the future. Most practitioners commented on the role of funding and how it influenced the project implementation, as Addison, practitioner, explained:

funding makes implementation very fragmented and makes it difficult to consolidate any approach. **If we have some more money to do this, let's do this...then money finishes so let's stop doing it.** (Addison, **emphasis added**)

Addison's depiction suggested that funding exclusively means money, because the outcome of any funding proposal was to receive, or not, money. However, there were other issues relevant for sustainable funding (such as competence, strategy, learning, and training) which seemed to be overlooked by NDEC. Addison appeared quite clear about the importance of funding by translating into what they do (or not do) in practice— *“we have money, let's do it; or the money has finished, let's stop doing it”*. In addition, another practitioner emphasised the size of the organisation, as Dale remarked:

for a small organisation like us, it can be difficult and in the end could just be the fact of the money, getting new projects that counts. (Dale, 2nd interview, **emphasis added**)

Dale illustrated how funding can act as a 'driver' not only for the project implementation and evaluation, but for the whole organisational strategy. Without funding the project and NDEC were unsustainable. Besides, the issue of funding was observed across various stages of the project and, particularly it was challenging to devise a strategy for NDEC in this unsustainable context, towards the end of the project, as Glenn, coordinator, illustrated:

At the moment the only thing helpful after April is **really money**, because **I may have vastly developed my evaluation skills and experiences but without any funded projects, those skills are useless**. (Glenn, 2nd interview, **emphasis added**)

The notion that the funding element shaped NDEC's evaluation approach also emerged from some practitioners' accounts, as Dale affirmed:

If you are relying on project funding you will go by funders' rules. (Dale, 1st interview)

Dale's affirmation, combined with other participants' articulations illustrated how the funding dependency influenced the way EP was designed, planned, implemented and reported. If the NDEC played by the funders' rules, their team should already know what to expect. From a coordinators' view, most funders included evaluation in their eligible costs and that appeared the single reason why these organisations evaluate their projects. This funding dependency sometimes seemed to play against them, as Eli, coordinator, reinforced:

Sometimes funders do a bit of a disservice to evaluation; they make demands that turn off people that deliver the evaluation work. (Eli, **emphasis added**)

Regarding the other coordinator, Glenn, emphasised the organisation's obligation to evaluate:

We have to conduct evaluation **to please and serve the funders; otherwise we would not have time do it.** (Glenn, 1st interview, **emphasis added**)

The perceived account of the practice of evaluation appeared to privilege funders' request as the main reason to do it. The context in which the YP evaluation was practised sheds light on the overlooked role of practitioners' competence. Moreover, the theme of (lack of) competence in evaluation was visible across participants' experiences of evaluation, particularly during the design and planning stages of evaluation. By having primarily focused on the element of funding, the role of participants' competence to evaluate seemed neglected, as Eli commented:

I am not involved personally **and I am not really sure where the evaluation is up to, I don't know** [*pause*] **I should know, but I don't know.** (Eli, **emphasis added**)

In this context of loose guidance, each challenge appeared massively complex and practitioners appeared to believe in their lack of experience to evaluate, as Sam, practitioner, explained:

In relation to evaluating and measuring the performance of this project, I think we have probably done half **and we are not quite good on that.** (Sam, **emphasis added**)

Some practitioners' accounts related to a clear priority given to money by disregarding the need for improving practitioners' competence to evaluate; whereas for others the idea was that they have not been competent enough in evaluation, as Stevie, practitioner, shared:

I am very much aware that with this project the monitoring and **evaluation that we do is a 'stone' thrown away from that ideal.** (Stevie, **emphasis added**)

The realisation that practitioners have to work hard to achieve competence in evaluation was reinforced by the notion that it was difficult, and almost impossible to evaluate. Most participants mentioned, during the first phase of the interviews, that they did not know how to evaluate because they lacked skills. However, from my longitudinal analysis it appeared that possible reasons for this perception were not only associated with the lack of skills, in terms of access to knowledge or expertise, but partly with internal and contextual barriers to practitioners' participation in evaluation (discussed in section 6.3).

Findings from the current study emphasised practitioners' belief in their inadequacy to evaluate, which corroborated Saunders' (2000) point that sometimes evaluation is not the central core of practitioners' WAs, which is certainly the case in NDEC; however, an inconsistency observed was that these

practitioners are asked to do it as if they were experts. As “*individuals working with a wide variety of professional preoccupations*” (*ibid. p.14*) they assumed and judged themselves as insufficiently skilled, thus they tend to resist in engaging with a performance-based evaluation (resistance further discussed in section 6.4.3). My study shed light on the preponderance of a funding-dependent organisation at the expense of their practitioners’ competence of evaluation. Consequently, they appeared to practise evaluation, with lacking confidence in their competence. Possible explanations for practitioners’ lack of confidence may relate to the barriers to participation, as explained next.

6.3 Barriers to practitioners’ participation

6.3.1 Leadership

In articulating their experiences of participation in evaluation, most practitioners shared that they had not participated in the project design, which may have led to the rejection of the project funding. Surprisingly for most practitioners, but not for NDEC’s coordinator, the project was approved and a variety of perspectives were expressed in this respect:

I inherited the project from Glenn, who **created it almost in their head.** (Sam, **emphasis added**)

The notion that practitioners did not participate in the project design was reinforced throughout my observations, particularly, when they were prompted about the possible reasons for their lack of involvement in participation. From the various reasons mentioned, the coordinator’s expertise in writing successful bids for funders stood out. Other reasons noted a multitude of other activities

that had to be dealt in a context of few project workers available. In elaborating about this period, Glenn, coordinator, acknowledged that the project had been designed in isolation without reference to others, which generated some misunderstanding about the evaluation and the project as a whole. Glenn retrospectively described their thinking at that moment of the project design:

It was partly difficult because **I** submitted the project **without particularly thinking about the outcomes**. When **I** submitted the bid, in all honesty, **I was planning** exciting activities. (Glenn, 2nd interview, **emphasis added**)

From Glenn's perspective, it was more about the excitement of doing something, rather than thinking about the implications for practitioners in the future, which include their lack of participation during the design stage, possibly influencing their practice of evaluation in the subsequent stages of evaluation.

The findings reported in this section highlighted that the role of coordinator as a leader seems a barrier to the participation of practitioners in evaluation. This study emphasises that although the practice of the leader has said to be mainly rooted in in the organisational context (Mintzberg, 2009), each leader's practice influences their team's perception and practice of evaluation. The practice-change episode (PCE) of practitioners' lack of participation in the YP's design illustrates that influence; particularly when the leader seemed confident in sole designed the project. Possibly from the leader's view they were having the best trade-off decision in terms of balancing the organisational context with internal management and financial sustainability; whereas their team' views thought otherwise – they appeared to perceive the role of leadership as a barrier to their participation in the project.

6.3.2 Funding

As the element of funding was so vividly expressed across my dataset, it interweaves the three main themes emerged from the analysis. For this reason, I refer to it various times from different perspectives; for instance, here I refer to funding as a barrier to practitioners' participation in evaluation, particularly in the stage of design, whereas in section 6.2.1, I referred to the funding dependency as a feature of the context.

Practitioners' articulations of the project design acknowledged that they have not been involved in the project designed by their coordinator, as reported in section 6.2.1. As a result of practitioners' lack of involvement, I observed a tension between practitioners and coordinators, because the former did not know the rationale of the project. During the planning of evaluation, the coordinator went on leave, one month before the beginning of the project, leaving their team solely in charge of the evaluation:

When I went on leave, the team had to develop an evaluation plan with the outcomes. **I think had I been here, it would have been more easily...probably it stretched them too far, and out of their comfort zone and experience** [*pause*] and generally we have not much time to evaluation. (Glenn, 2nd interview, **emphasis added**)

When asked to describe this moment, practitioners explained that the funding had a major role in the decisions around practice - design, planning, implementation and reporting - of evaluation. For example, Jesse, practitioner, shared that:

it was **only once we received the funding that we became engaged** and involved in the evaluation. (Jesse, 1st interview, **emphasis added**)

This example illustrates that funding can act as a barrier to practitioners' participation, as well as shapes when and how they become engaged in the evaluation. This study reinforces that the barrier of funding influences the participation of practitioners across the YPET, especially during its design, because it appears conditional upon the status of funding receipt. This explanation challenges Stake's (2004) view of practitioners (and others stakeholders) as active and engaged collaborators. In my case study, they have to overcome some barriers in order to fully engage (Wenger, 2008) in EP and in the process their perception of evaluation is being shaped. DSP literature assumes that the meaning of the practice of evaluation is depicted by the social and symbolic consequence of participation each moment over time (Shove *et al.*, 2012). My study adds that the lack of participation has also consequences, especially in terms of practitioners' engagement in the later stages of evaluation.

The element of funding was seen as a barrier to participation, observed in all stages of the YPET, which adds to the EP literature (Simons and McCormack, 2007) that advocates an increased participation of stakeholders in evaluation, by proposing that funding, as a barrier to participation, needs to be considered.

6.3.3 Co-occurrence of working activities

My analysis suggested another perspective for the lack of practitioners' participation in the project design, which was the existence of concurrent working activities (WAs) in the organisation. Arguably, it seemed that the

coordinator acknowledged NDEC's lack of time and resources by privileging the organisation's efficacy, in terms of their ability to produce a desired or intended result, rather than involving practitioners in the project design, which inevitably concurred with their other WAs. From the practitioner's perspective that co-occurrence seems part of their daily routine, as Jesse suggested:

lots of other things in the organisation **are happening at the same time**. (Jesse, 2nd interview, **emphasis added**)

The notion of co-occurrence seemed to have worsened when the coordinator was on leave and Sam, practitioner, had to respond to GMA's requests, as explained:

We documented it all which is one of the things that GMA tells you to do is to document every stage and any incidences that you have been made. (Sam)

The episode of the coordinator's absence was interpreted as a PCE (see Table 6.1; PCEs discussed), because it prompted change to occur in practitioners' EP. Whilst concurrent WAs appeared to explain, to some extent, why practitioners had not participated in the project design, their 'forced participation' (discussed in section 6.4.2) led to the co-occurrence of WAs with the YPEP across timeline (Figure 6.1).

Previous studies on NPOs' EP have been inconclusive regarding the connection between stakeholders' participation and the co-occurrence with other WAs (Carman, 2007; 2009). As these studies were conducted in medium or large NPOs (Carman, 2007) within institutionalised contexts, as universities or European partnerships (Saunders, 2000; Saunders *et al.*, 2011), the co-occurrence of WAs may have not been raised, because possibly these large

organisations have a dedicated unit for evaluation, so that practitioners have not experienced concurrence in those settings. However, participants' experiences suggested that in a small DE organisation, such as NDEC, with unstable human and financial resources alongside a lack of sustainability, they engage in multiple WAs such as, management, implementation, and monitoring (further analysed in chapter 8). This finding allowed me to infer that practitioners had not participated in the project design, possibly, because they were already involved in concurrent WAs within NDEC. As an illustration, NDEC, as an organisation, appeared to not involve practitioners in the design, because they were also already participating in other concurrent activities, as Jesse illustrated "*other things are happening at the same time*". This inference also raised the possibility that their coordinator had experienced the pressure of co-occurrence of WAs before, and wanted to protect their team from the same experience; thus evidence seems inconclusive about this point. In this vein, the co-occurrence of WAs represents a barrier to practitioners' participation in the project design, which influences their participation later in the evaluation.

In addition, my study extends Carman's (2007) work on how evaluation is practised in NPOs, and in the specific domain of DE, by pointing that participation of practitioners in evaluation can be influenced by the co-occurrence of other WAs. This co-occurrence may threaten their deliberate participation in the design and EP. Consequently, the participation of practitioners tends to be dictated, managed and regulated by others leading to a sense of lack of ownership of EP (further discussed in section 7.5.3).

In summary, the combination of these accounts suggested another layer of complexity in the participation of practitioners in evaluation, visible through the three barriers discussed. Within this context of a convoluted participation across YP evaluation, a non-linear sequence of the typology of participants' engagement is discussed next.

6.4 The typology of participants' engagement in the practice of evaluation

6.4.1 Disengagement

In reporting a PCE of the project design, I refer to practitioners and coordinators as 'participants'. The project design stage sets the scene for accommodating (or not) the EP, particularly, when NDEC was asked from the beginning to reflect on the planning of the YP evaluation. Surprisingly, most practitioners indicated they have not participated in the project design and so had difficulties in planning the evaluation of an unknown project, as Dale, practitioner, mentioned:

I haven't been involved in anything with DfID. (Dale, 1st interview)

The design stage was observed by participants as fundamental in the development of the project and when asked about what led NDEC into that stage, Jesse, practitioner, recalled:

The project had been designed by our coordinator, so **I wasn't involved in the design of it at that stage. It had been informal conversations but not any formal conversation in the sense, let's think this out.** (Jesse, 1st interview, **emphasis added**)

Jesse's view was echoed by most practitioners emphasising their lack of participation in the project design. The project's single-led design by NDEC's coordinator was one thread cutting across the subsequent stages of the project; as it seemed to have influenced the practice of evaluation, as Dale, another practitioner, illustrated:

from my point of view it would had been better for our coordinator, to have communicated about how are we going to achieve the outcomes that we laid down [*pause*] **because they were self**, we were the ones actually saying that we will achieve those. (Dale, 1st interview, **emphasis added**)

A recurrent issue across the two phases of interviews was the disengagement between the project design and its evaluation, particularly in terms of the nature of targets written in the proposal. Glenn, coordinator, shared that the project outcomes had not been thought about, in terms of evaluative thinking and that seemed to complicate the current planning of evaluation. As a result, in the later stages of implementing and reporting evaluation, practitioners mentioned they had difficulties in conducting the evaluation (possibly, because they did not know what Glenn intended when the project was first designed). The perception emerging within the team was that:

evaluation is not a practice that anyone could do and certainly is an impossibility to do all that evaluation entails, and to do it well. (Jesse, 1st interview, **emphasis added**)

In prompting funders' views about how DE organisations engage in evaluation, Kelly, manager, explained:

A lot of organisations really don't have a good grasp between doing one activity and trying to work out what the impact of that activity is; the need for understanding what changes they want to

make and how future activities should draw on their learning to bring about the changes they want to achieve. (Kelly)

Regarding the connection between design and evaluation stage, the findings reported suggested that NDEC is yet to design their interventions, according to funders' guidelines; but rather according to their own resources, competences and motivations. This inference may be transferable for other NPOs alike (see section 9.4; implications for future research). Although some M&E handbooks place evaluation at the end of the project cycle (EuropeAID, 2004), some funders position it at the outset, even requesting evaluation plans at an early stage (DfID, 2003; Renz and Herman, 2010). As a result, the design, and subsequently evaluation has constantly been changing, depending on the resources available. Whilst previous literature acknowledged some developments in re-positioning the evaluative thinking (Saunders *et al.*, 2011) from the start of the project, my findings highlighted the perception of evaluation as a non-priority practice. Moreover, engagement is interpreted as a form of belonging and identification with a community of (evaluation) practice (Wenger *et al.*, 2015). My study sheds light on a disengaged moved that emerged from practitioners' perception of design and evaluation as dissociated stages.

This study also demonstrates that the disengagement between design and evaluation generated an inconsistent level of participation of practitioners in EP. For example, the thinking and practice of evaluation may have emerged later than expected in the YPET possibly because of the barriers to participation, discussed in section 6.3. The study highlights the possibility that NDEC still perceived evaluation as the final stage of the project cycle management (PCM),

which clashed with funders' increasing expectation of establishing evaluative thinking from the start (DfID, 2015). My research illustrates the influence of the participants' misperception of the PCM in not keeping the pace regarding the connection of design and evaluation of the projects. An implication of this illustration and main finding was that the design of the YP project appeared influential to the way evaluation is practised.

6.4.2 Forcing

From the aforementioned description of disengagement in the project design, practitioners seemed to not expect a necessary involvement in the evaluation, during its planning stage; however, as their coordinator went on leave, they were asked to step in and take the lead. When prompted about this period, Glenn noted:

I was on leave so I wasn't here and from all accounts it has been a **pretty nightmare process**. (Glenn, 1st interview, **emphasis added**)

Their sense of the team's lack of competence, time and expertise suggested an unintentional engagement in evaluation, as Jesse, practitioner, described:

We have only been evaluating as it is forced upon us. (Jesse, 1st interview, **emphasis added**)

There was some sort of 'forced' action upon NDEC, which appeared to have prompted practitioners to engage in the evaluation, even without a previous involvement in the project design. This forced engagement in planning the evaluation highlighted an adaption, from the previous disengagement in the design, to an opposite extreme of a forced engagement in the subsequent stages

of the evaluation – an adapted practice of evaluation (further discussed in section 6.5).

The findings, reported thus far, revealed that NDEC practitioners neither participated in the design nor were involved in the YP decision-making process (Greene, 1988) as engaged collaborators (Stake, 2004); or even asked to bring their commitment, resources and skills (McCluskey, 2011) to their practice of evaluation. Subsequently, my study indicated that the engagement of practitioners in EP may not occur; or may occur without their previous participation in the design of the project as a perceived forced practice of evaluation.

Literature seemed to take-for-granted the participation of practitioners (Greene, 1988) across all stages of the project cycle; yet my study indicated that when practitioners had not participate in the design, they were likely to struggle with a ‘forced’ participation in the planning of evaluation later. This perceived struggle seems reasonable, as practitioners did not know what has been designed and planned to achieve, therefore they struggled to figure out what and how to plan the project evaluation. This study challenges the literature assuming the participation of practitioners from the beginning of the evaluation process, by emphasising that their active participation may not always follow a conventional route of engagement, as happened in the NDEC case. Within this route, barriers to leadership, funding, and co-occurrence of WAs need to be overcome, to engage practitioners in EP as they resist doing so, as resistance lies within their selves. (See Figure 6.1, timeline of PAR-YPET).

6.4.3 Resistance

In planning evaluation without coordination, practitioners shared their perceived difficulties, as Sam recalled:

[The] coordinator left in March and the project started in April so I was in the deep end with the new DFID's funded project with some really heavy evaluation tools that GMA wanted **and I did not have that much experience in evaluation and monitoring.** (Sam, **emphasis added**)

Most practitioners acknowledged that they had to incorporate their coordinator's tasks within their own daily practice, even the tasks requiring more expertise and responsibility, and at this stage, both practitioners' interviews and my field notes suggested a certain level of resistance: they wished they had their coordinator there for guidance. Also, another recurrent observation during the planning stage concerned the performance of evaluation, as Jesse explained:

So if the evaluation relates to performance in any way it won't do, it will not be reflective, people will tell you just what they think you want to know. (Jesse, 1st interview, **emphasis added**)

Other practitioner perceived this stage of evaluation as "*a huge task dictated from above*" (Dale).

During the evaluation reporting stage, the PCE of a film evaluation was observed, when Dale inquired the funders about other informal methods of reporting (for example, a film evaluation), as outlined:

While film and photos were very useful I am not sure how much the funders[*pause*] still very much about the written reports, **so it is only in addition to, rather than, certainly not substituting.** (Dale, 1st interview, **emphasis added**)

Funders ended up reinforcing the need of a formal evaluation report, rather than a film evaluation, and Dale and other colleagues not only believed they had been forced to evaluate, but also perceived the power of funder's decision. Despite some resistance from practitioners (and also from funders), the YP evaluation was implemented through a formal report delivery.

As discussed in section 6.4.1, the disengagement in project design may explain practitioners' resistance in participating in the evaluation. Whilst practitioners seemed reluctant to evaluate, resistance was also evident with funders' unwillingness to change their approach to evaluation reporting. The dominant element that resists change is inmost. Previous literature on resistance appeared to be mainly unidirectional, i.e. focused in one single direction of resistance from practitioners towards funders (Taut and Brauns, 2003). However, although my study recognised the dominance of this unidirectional nature of resistance, it also identified a bi-directional resistance from funders towards practitioners (as they resisted changing the evaluation method to a more informal film evaluation). My study extends that of Taut and Braun's (2003), by suggesting a bidirectional dimension of resistance when engaging with evaluation (practitioners- funders and funders-practitioners). My study also sheds light on new interpretations of evidence when reporting evaluation such as arts-based methods, because their outcomes have more power to induce responsiveness than most methods of evaluation, principally reports. Nevertheless, my study adds that despite the attempts to use some of these methods through a film evaluation proposal, the dominant voice was still the funder's; which might have increased practitioners' frustration.

Literature on EP has focused on the ontological nature of resistance (Saunders, 2000), leaving unattended the analysis of resistance as a problematic obstacle to the practice of evaluation (Datta, 2001). It has been argued that practitioners resist in engaging with evaluation by emphasising a *defensive* approach or by avoiding participation in evaluation (Saunders et al., 2011). My study proposed a strong connection between resistance and the issue of participation, particularly regarding practitioners' deliberate exclusion in the project design. I argue that practitioners' resistance towards evaluation may compromise their practice of evaluation over time, because it generates indifference and passivity, leading to a lack of participation. In recognising participants' difficulty in acknowledging evaluation as a daily practice, my study corroborates Datta's (2001) analysis of resistance as a problem that may jeopardise evaluation practice over time.

Previous literature on evaluation interpreted resistance as a consequence of power relationships, specifically, of the attempts made to maintain a certain power or status quo (Taut and Brauns, 2003). My study adds a novel insight to the literature, by proposing that resistance in evaluation can be perceived as an obstacle to participation in evaluation; i.e., in a DE context, a bi-directional resistance was also impeding the course of practitioners' engagement in evaluation. This was observed, for example, when practitioners were supposed to participate in the evaluation, having not been previously participated in the design of the YP, generating resistance (see section 6.4.1).

The current study suggests that through an unconscious practice of evaluation, practitioners sometimes cannot recognise their resistance, because what one

perceives doing, is different from what they actually do (Shove *et al.*, 2012); which was illustrated in the PCE about the film evaluation; or when Glenn resisted in involving their team in the design, which was perceived by Glenn as acceptable, given the NDEC's circumstances. This feature of an unconscious resistance when participating in the evaluation emerged from a SPT consideration of the role of resistance in the formation and development of practices over time (see chapter four). This feature extends Saunders *et al.*'s (2011) acknowledgement of the mere existence of resistance in EP. In addition, a fresh insight from my study illustrated how practitioners' resistance, in engaging with EP, has changed across the YPET from a disengaged and forced pattern in constant adaptation towards a coping and collaborative practice, discussed in the next two sections.

6.4.4 Coping

The sharing of difficulties and uncertainties amongst practitioners was perceived as a critical mechanism to build the project EP. A sense of leaping into an unfamiliar territory emerged, during the evaluation planning stage, yet, practitioners had to manage a three-year demonstration of hard evidence, plus planning, implementing, and reporting a prescribed evaluation. When attempting to reflect about this period, Tyler, practitioner emphasised:

I have not really thought about that [evaluation]. I am here to do what needs to be done. (Tyler, **emphasis added**)

Tyler seemed to neglect thinking over practising, possibly, because they had so many other concurrent activities to respond that they could not afford to stop

and reflect about what they were really doing (see section 6.3.3 on co-occurrence of WAs; also discussed at length in chapter 8).

As the team was facing new challenges, I observed their effort in moving ahead with the YP evaluation, regardless the constraints. On the funder's side and despite the managerial discourse, they made an effort to clarify what has been perceived as a stricter evaluation approach, by appointing a 'grant management agency' (GMA) to monitor and evaluate the project. In those circumstances, GMA encouraged NDEC's practice of evaluation, as Glenn, coordinator, emphasised:

GMA told to all DE movement: "Look, **you've got to improve your evaluation**". (Glenn, 1st interview, **emphasis added**)

In recognising the need for support, the team appeared resilient, accepted the tensions encountered and moved on towards a tough period of hard work. This recognition was particularly relevant during the stages of evaluation planning and implementation, when tensions about performance as the final purpose of the evaluation, arose; as noted by Jesse:

So if in any way evaluation relates to performance, it won't happen! (Jesse, 1st interview)

Another practitioner added:

One thing that takes a lot of time and it is difficult in our work is trying to evaluate [*pause*] you know [*pause*] how you reach those targets. As soon as you put in the numbers games **and the need to tick boxes it becomes very difficult to enjoy it.** (Dale, 1st interview, **emphasis added**)

Curiously, there were very few occasions when terms such as ‘*enjoyment*’, and others alike, came through the dataset, which may suggest a negative experience of EP and/or the inadequate evaluation approach currently requested by DE funders. During these stages, practitioners coped with conflicting views (e.g. whether evaluation was only measuring YP performance) and unfamiliar issues, moving out of their comfort zone. Coping with difficulties and uncertainties was critical to move out from a resistant default position in evaluation practice (see Figure 6.2, PAR-YPET).

Coping, as a form of engagement, led to the beginning of practitioners’ process of participation, with a ‘forced’ nuance, though. In coping with difficulty, practitioners created a sort of willpower to accept that they had to improve their evaluation, as requested by funders. Through coping, they believed to own their practice of evaluation, which was illustrated through experiences such as, a dictated practice of evaluation or a contest against an isolated practice of evaluation. In coping with these difficulties, practitioners were encouraged to team up to implement the evaluation and these strategies were interpreted as coping strategies that catalysed the change in their engagement. These strategies emphasised the centrality of their willingness to co-operate, rather than perpetuating their disengagement.

This study expands literature on evaluation informed by CoP (Saunders *et al.*, 2011; Wenger *et al.*, 2015) by proposing a coping dimension to deepen the understanding of engagement in evaluation. Resilient practitioners coped with toughness even in situations where they had not realised so, as Tyler’s example illustrated. As such, my study highlights that engagement occurs even without initial belonging or identification (Wenger *et al.*, 2015), but rather within a non-

linear pattern of disengagement and resistance. Through coping, a community of (evaluation) practice can be reconfigured and changed, within a DE organisation. Moreover, a sequence of cycles of coping results in a progressive change towards collaboration in EP.

6.4.5 Collaboration

The YP reporting stage was critical to align expectations, to clarify misunderstandings and to prepare for the external evaluator's support. Some observations of this period were captured in the Extract 6.1, below:

Where is the Summer mood?

It is August, and the office is almost empty. One or two practitioners wander around...they seem always exhausted and demotivated. They have now finished the project's second year report and they received the confirmation for another year of funding. That will be the last year of implementation. They learnt...they overcame barriers and faced challenges, but somehow it seemed that something was missing; that their effort was never good enough, never fulfilling them. Dale is worried with getting the right evidence to demonstrate the project's achievements. Perhaps, that could be one of the pieces missing. They seemed all so frustrated and stuck.

August 2012

Extract 6.1: Extract from field notes, August 2012

Practitioners combined efforts to work towards the delivery of the YP evaluation, which may have been prompted by the realisation that they need to do it in order to get further funding. In this line, I invited funders to describe how they assessed DE evaluations, which Kelly, manager, described:

Certain decisions are done by me or my senior manager in terms of recommendation for funding, but we pay GMA to manage the DAF fund on our behalf. Each year we have a **competitive**

funding round and GMA presents for us the results of the candidates and its scores. (Kelly, emphasis added)

Kelly went on outlining their view on whether DE organisations should have an effective results framework to measure change:

I think the main aspect is really making sure that each grant holder has appropriate and effective results framework, which really measures what is the change that the project is trying to bring about and then how do they measure whether that change has happened. (Kelly)

As Kelly shared these expectations with the team, through GMA, the effect was more than an initial fear, a collective call for support, responded by the appointment of an external evaluator, Brook. The observations from this period of implementation allowed me to depict that both coordinators have relied on Brook's expertise to improve the evaluation, and evidence about their role, was in a confidential appendix to ensure their anonymity is protected.

The process of sharing difficulties, by working collaboratively, seemed to have gradually engaged practitioners in recovering their sense of ownership, during the implementation stage. In this vein, this study found that a collaborative engagement occurred after a moment of 'reality-check', illustrated by the funders' views on what was really expected from the NDEC team. A collaborative practice of evaluation was adapted from a period of coping and change on practitioners' engagement in evaluation was observed over time (see timeline PAR-YPET). The study also revealed that a collaborative engagement emerged as last in the following sequential order, after an instance of disengagement, forcing, resistance, and coping – a non-linear typology of engagement.

Previous literature on collaborative evaluation (CE) (Berner and Bronson, 2005) assumed that most stakeholders participated in all stages of the evaluation and in the decision making-process. The findings reported on the lack of participation indicated that in a specific DE project, stakeholders have an inconsistent level of participation in evaluation and their engagement may vary in form over time. My study illustrated that a CE was only pursued towards the end of the project, due to the specific constraints of the context, namely, the absence of NDEC's coordinator, the uncertain future of the organisation and the need of additional funding after YP completion. Despite this adversarial environment, these findings indicated the realisation from practitioners that they had no other solution than, reluctantly or not, accept the 'real-time' constraints to move towards a necessary collaborative practice, supported by the external evaluator.

Thus, this study contrasts with those of Berner and Bronson (2005) by emphasising a dynamic pathway in the collaborative engagement in evaluation, from a 'forced' to a 'deliberate' participation. Also, it challenges the notion of a deliberate collaboration by adding the dimension of co-construction. So, the collaboration observed had an implicit tone of survival and it seemed almost coercive sometimes, but despite this pressure, practitioners ended up collaboratively co-creating their version of collaboration, which meant that practitioners, consciously or unconsciously, have accepted instances of change in their context, as well as instances of change in their engagement in evaluation. It also meant that a CE has possibly been pressured by the research process undertaken, and in this sense co-constructed with my presence, as a researcher. Furthermore, a CE has been co-constructed with the wider community, for example through the external pressure of funders and external

evaluator. My study adds to the literature on CE (Berner and Bronson, 2005) by proposing that collaboration can be constructed through a non-linear typology of engagement of practitioners in EP, changing over time. Collaboration may have not occurred as a CE suggested, in terms of practitioners' participation in all stages of the evaluation and in the decision making-process; but in a DE domain it still occurs under a different meaning, under a social practice view- the need of delivering a very tangible and prescribed evaluation report by March 2013, in order to guarantee the organisation's financial sustainability¹³ for next year, rather than closing down and make all team redundant. A SPT view of evaluation accepts this co-construction of collaboration, as the meaning of the practice of evaluation is depicted by the "*social and symbolic significance of participation at any one moment*" (Shove *et al.*, 2012, p.23). At that moment in the YPET the significance of participation was through a co-constructed collaboration.

Other aspect relevant to this discussion is the type of decisions in which stakeholders were likely to have an input, such as whether and what to evaluate, how to elaborate conclusions, when to circulate results, and how and when to implement further recommendations (Worthen *et al.*, 2004). These decisions are likely to inform and be informed by stakeholders' practices over time, as this thesis argues throughout. As noted, and in contrast to CE literature, practitioners and coordinators did not participate in the decision-making process of evaluation, because most decisions had already been taken by the funders. Other

¹³I am not convinced that funding grants are a financially sustainable source, but that is another discussion.

possible reason is that stakeholders' opposing views about evaluation threatened a collaborative engagement in evaluation. This finding corroborated the literature on program evaluation, advocating that the stakeholders' opposing views on evaluation influenced their practice over time (Saunders *et al.*, 2011).

In summary, my study contrasts with that of Berner and Bronson's (2005), particularly regarding the inclusion of practitioners in all stages of evaluation. This contrast could have occurred, because of the type of evaluation practitioners were engaged with, was not a collaborative evaluation. Practitioners deliberately engage to collaborate in the evaluation, but the meaning assigned to collaboration could have been expressed by different mental constructs, relationships and perceptions (Shove *et al.*, 2012).

6.5 The change observed in evaluation practice: the paradox of participation

In relation to the theme of participation, the change observed in EP was through adaptation over time (see Figure 6.2; PAR-YPET). The effect of modification and adaptation appeared almost extreme and I termed it the 'paradox of participation', because practitioners' participation appeared at first disengaged, then adapting to a forced practice, through to a resistant position that was then integrated through coping strategies towards collaboration. The paradox observed was a modification and adaptation from a lack of participation in the design to a collaborative practice of evaluation in the last stage of reporting. Such participation occurs within "*a landscape of possibilities that is, in any case, always in transition*" (Shove *et al.*, 2012, p.145).

6.6 Summary and conclusions to the chapter

This chapter reported and analysed the theme of participation in the YPEP. Findings and discussion spanned across four sections as summarised in Table 6.3 (overleaf).

Participation (PAR)	Findings emerged	Location
The context of the YP: disconnect of design and evaluation	1) The design of the YP's project is influential to the way evaluation is practised 2) Design seen as 1st phase of project cycle and evaluation as the last phase.	Section 6.2
Funding-dependency vs competence in a DE working environment	A funding-dependent working environment jeopardised the focus on practitioners' competence of evaluation.	Section 6.2.1
Barriers to practitioners' participation	1) Leadership: unilateral decision of a coordinator to solely design the project 2) Funding: once received, gateway for practitioners' overall participation in the project. 3) Co-occurrence of practices: practitioners had other multiple working activities at the same time, which undermined their participation in the project's design	Section 6.3
The engagement of participants in the practice of evaluation	A typology of engagement was observed, from disengagement, forcing, resistance, coping and collaboration in EP.	Section 6.4
The change observed in EP: the paradox of participation	Practitioners disengaged in the project' design, but forced to participate in the evaluation.	Section 6.5

Table 6.3: Summary of findings on participation (PAR)

The chapter concluded that as the context in which the YP was evaluated is funding dependent, funding was prioritised over practitioners' competence. It emphasises three observed barriers to practitioners' participation in the project design - leadership, funding, and co-occurrence of working activities - which I argue are likely to influence the EP at subsequent stages.

The chapter also revealed an inconsistent level of participation of practitioners in YP's EP, illustrated by the disengagement between a project's design and its evaluation. It proposed that this disengagement of participants in evaluation jeopardises their participation over time. Hence, funders' request for evaluative thinking from the beginning of the project is yet to be applied in practitioners' minds.

The chapter identifies a typology of engagement of practitioners in evaluation through disengagement; forcing; resistance; coping, and collaboration, which influenced their trajectories of participation across the YPET. The chapter discussed participation from a DSP lens, which meant that practitioners are understood as carriers of the EP. The way they understand, desire, avoid or want, are features of their singular participation in EP, and not qualities of their individuality. In this chapter, I argued that practitioners are yet to fully participate in all cycles of project management, and also that the issue of participation influences how they practise evaluation over time.

7 POWER RELATIONSHIPS IN THE EVALUATION PRACTICE

7.1 Introduction to the chapter

The chapter starts with a critical account of the context in which power relations (PR) emerged and were perceived by participants (7.2). Section 7.3 describes and discusses three configurations of PR observed across the project timeline: a leadership-led PR, an expected-led PR, and a competence-led PR (CLPR). Section 7.4 focuses on the influence of PR in the practice of evaluation regarding the role of perception – three episodes of influence were observed between funders and coordinators, coordinators and practitioners, and amongst practitioners. Section 7.5 explores participants' accounts of the consequences of PR for the practice of evaluation, in a DE context, using three practice-change episodes (PCEs) illustrating a legitimised dishonesty, the role of judgement in evaluation, and the influence of scepticism. Section 7.6 describes the change observed in the EP, partly propelled by the influence of PR over time. Section 7.7 summarises the findings and conclusions of the chapter. Throughout these sections, the chapter discusses the theme of PR comparing findings with those in the literature reviewed in chapter two.

The chapter begins with a visualisation of the issue of PR and its emergence across the youth project evaluation timeline (PR-YPET). The rest of the chapter unpacks Figure 7.1 (overleaf).

FINDING Power relationships (PR)

FRBC Funder-recipient based context (circular dynamic of change – modification, stabilisation, dominance)

CPR Configurations of power relationships

CPR-EP Consequences of power relationships for evaluation practice

C-EP Change observed in evaluation practice: by modification and competition

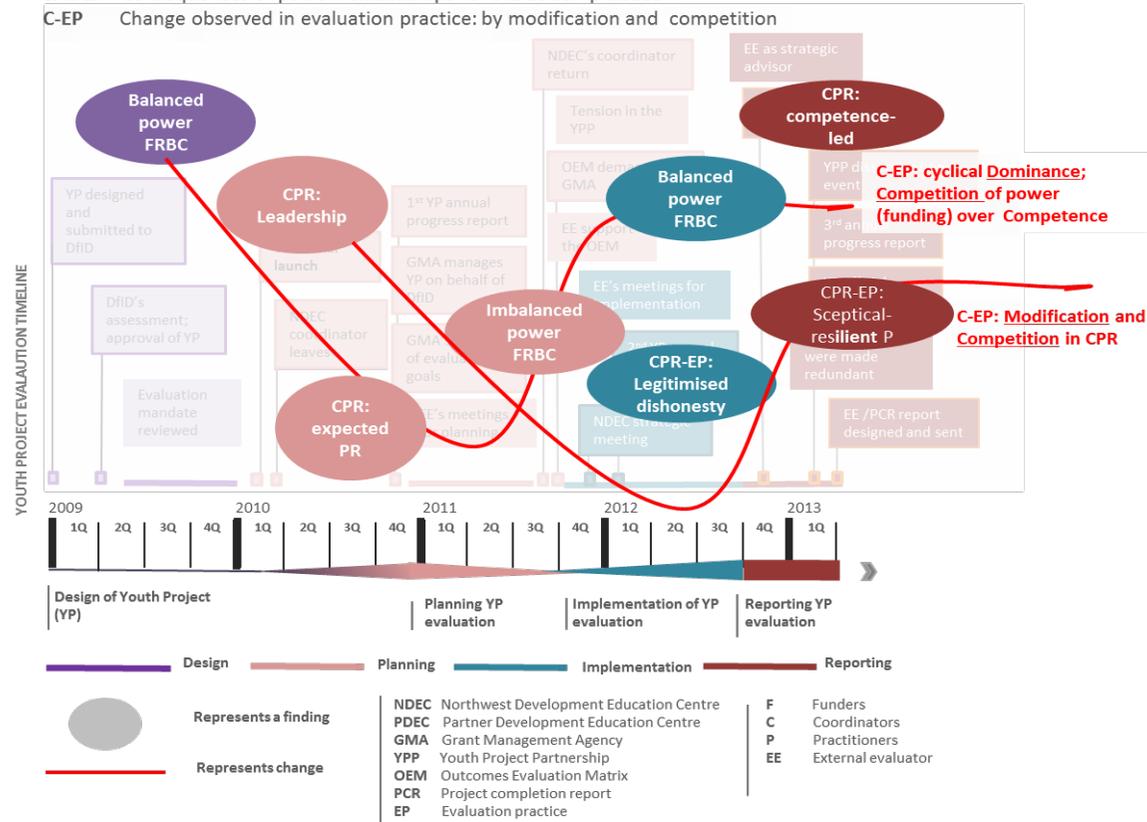


Figure 7.1: Power relationships in the youth project evaluation timeline (PR-YPET)

7.2 The context of a funder-recipient based evaluation: balanced and imbalanced-power relationship

When describing the context in which the youth project (YP) was designed and submitted to DfID, most practitioners reported either directly or indirectly the issue of power as in the following examples, respectively:

They [*DfID*] have the **power** just to pull the funding. (Sam, **emphasis added**)

We [*NDEC*] have to evaluate because **it is required by funders. Had they not required that, we would have not evaluated it** [*YP*]. (Addison, **emphasis added**)

One manifestation of PR observed by practitioners and coordinators was that relationships with funders had always been characterised by power. For example, Eli, coordinator, shared the view of a pressured practice of evaluation by funders:

People sometimes see the evaluation as something that the funder wants [*pause*] **so I have to do it for the funder, rather than what can we learn from that process.** (Eli, **emphasis added**)

During the evaluation planning, most participants shared Eli's view of the funder as powerful and when asked about their routines of evaluation, practitioners mentioned that they had managed other projects funded by DfID. This was the context when NDEC and their partner were informed that the project had been provisionally accepted. The partnership was also told that a grant management agency (GMA) would contact them to explain the details of the contract. A funder-recipient based context was visible as the main setting in which the YP evaluation was about to be practised, so issues of power appeared

prominent. From the beginning, a sense of insecurity on how to implement and evaluate the project was shared by both organisations.

I think one of the key things about evaluation was that nobody in NDEC had much experience and I think it is true to also say that **most of [*the*] evaluations that have been taken (*sic*) place in terms of past projects were a report, rather than an evaluation.** (Jesse, 1st interview, **emphasis added**)

I don't think that some of the activities delivered were very successful because we're actually targeting schools, rather than the non-formal side of things. (Eli)

This atmosphere may have been related to the Government's review of the funder's mandate for DE evaluation, or to how partners perceived their competence in evaluation (see section 6.2.1; funding dependency versus competence). Whilst participants said that funders previously had clear rules when they evaluated mini-grants, a redefinition of funders' goals was observed, regarding their approach to evaluation, which seemed to have altered NDEC's practice of evaluation. That alteration was manifested when GMA assumed the responsibility for monitoring and evaluation (M&E), subsequently altering the PR amongst participants. These depictions characterised this as a funder-recipient context, in which a powerful funder was represented, first, by DfID (design stage) and then by GMA (during the subsequent stages of evaluation); whilst the powerless recipient was represented by the project partnership. While the funder's power was attached to the funding element, where was the YP partnership power? Why did they seem powerless? Tyler and Jesse analysed how funders were perceived:

This funder [*GMA*] only administered the grant, they weren't development education specialists. It's like asking your bank what

are the best sausages, you know...they don't know [*laughs*]. They might eat it, but they don't know. (Tyler)

If I have a new colleague with experience in the evaluation I will say to them, can you think of ways in which we can evaluate, **but not in the funder's mechanical sort of way, not in a form filling sort of way, but something which has credibility**, because **I am not convinced that the way we evaluate this project has any real merits**, that actually was **very worthwhile**. (Jesse, 1st interview, **emphasis added**)

Interviews and observations suggested that participants perceived funders as having power which influenced how they practised evaluation across the YPET. Nevertheless, my field notes allowed me to infer that the way a powerful funder expected evaluation to be perceived, might be different within NDEC.

Manifestations of meaning in daily life

Today at NDEC's office I was paying attention to hear, see, and read anything with the word evaluation in the expectation of observing something explicit like the project folders, labels or some manuals in the bookshelf. I found one electronic file named "Monitoring and Evaluation" I navigated through a ton of documents where the word evaluation was non-existent – despite the file's name. I was puzzled! I came home disappointed and thought: so, what do they really mean by evaluation?

May 2012

Extract 7.1: Extract from field notes, May 2012

For example, this field note illustrated that M&E was conducted mainly because of funders' demand, and that is not always visible at an internal organisational level (e.g. in the content of the organisation's project folders). This study highlighted that PR were a pivotal element to the way participants described the context of evaluation, because they frequently articulated their ideas against the explicit power of funders. For instance, during the project's approval, participants experienced a balanced-power, which meant that in a funder-recipient based context, the rules were clear: funders' requirements were explicitly shared, so that coordinators were expected to respond to those

guidelines, knowing in advance their duties (e.g. the number of financial and technical reports and their respective deadlines), and assuming that they knew how to respond accordingly (Carman, 2007). My study suggested that funders expected to have their requests responded to, and coordinators expected to respond to those requests. The intersection of these expectations seemed powerful.

Conversely, participants' articulations, during the evaluation planning stage, revealed an imbalanced-power. GMA was the focal point for evaluation in the already convoluted funder-recipient based context. This generated an unequal division of power - funders seemed to have increased their powerful voice, by re-arranging and dictating how evaluation should be practised, by when and by whom; whereas coordinators and practitioners appeared to have a powerless voice, by losing ownership of their work. The observed shift from a balanced to an imbalanced-power relationship may partly be explained by the previous change in funders' expectations for evaluation, or by a misunderstanding by coordinators and practitioners about the review of the funders' revised expectations.

I am using the terms 'shift' and 'change' purposively to distinguish a rapid change over a short period of time (shift), from a change that implies a longer period (change). As such, previous change on funders' expectation may have shaped the shift from a balanced to an imbalanced context; that occurs and re-occurs in a circular dynamic. Figure 7.2 (overleaf) visualises the change of PR, observed in a DE evaluative context.

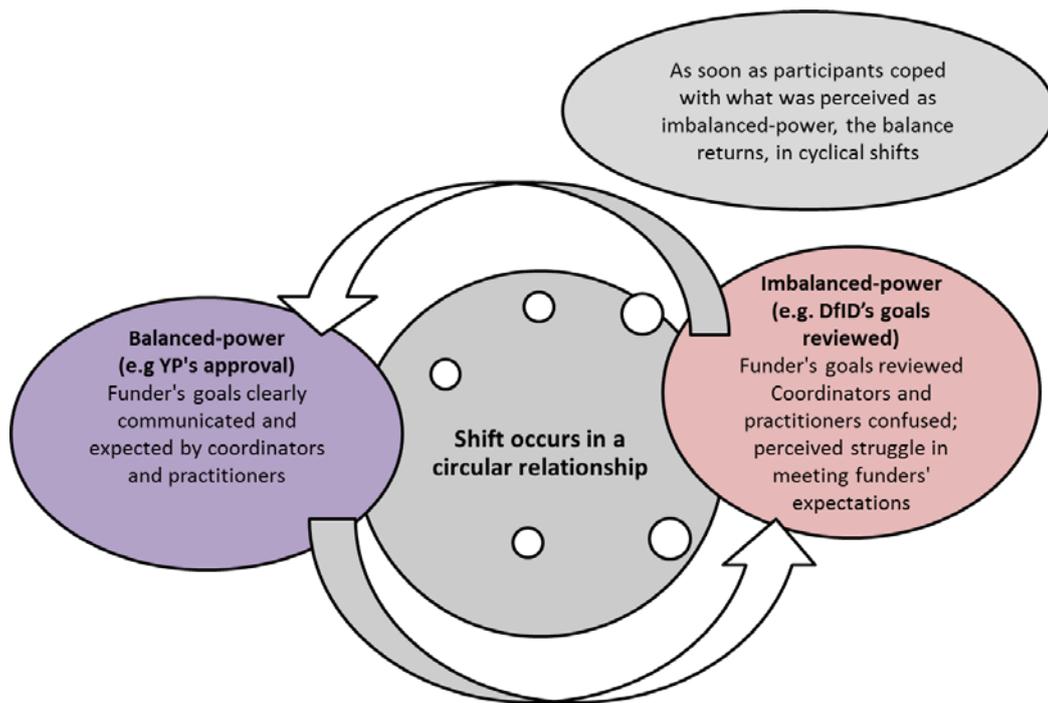


Figure 7.2: Balanced and imbalanced power-across the YPET

That change seemed to have occurred in a circular dynamic in the PR, starting with the shift from a balance (design stage) to an imbalance of power (planning stage) visualised by the bottom arrow in Figure 7.2. A re-occurrence of this cyclical PR was observed over time, for example, what has been seen as an imbalance of power appeared to have reached a sense of (re) balance again, during the later stages of the project, visualised by the top arrow. Once an imbalance of power had been experienced by participants, it might be perceived as balanced again over time; so changes in the balance of power could be perceived by participants over time (see section 6.4; typology of participants' engagement in the practice of evaluation).

The sphere of decision-making has always been where the money lies, because it represents the power of control between unequal groups (Saunders *et al.*, 2011). Funding or money represents the locus of control from funders towards

recipients; from the powerful towards the powerless. This was especially relevant within a funder-recipient context, where a powerful funder controlled the money and a powerless recipient needed that money. It may be that the inequality was driven by the element of funding. The findings from the current study are consistent with those of Saunders *et al.* (2011) in illustrating participants' expressions of power as a key influence to the practice of evaluation; but with the added insight that the balance of power may change over time, as visualised in Figure 7.2; balanced and imbalanced power across YPET, and further discussed in section 7.6; the change observed in EP. A cyclical and dynamic change of PR appeared to influence every single practice and activity, which expands the literature assuming a trend of a power imbalanced context in complex settings (Westley *et al.*, 2006). Therefore, the way PR influence the practice of evaluation also changes accordingly.

In addition, this thesis indicates that power can be exerted through meaning (Shove *et al.*, 2012), expressed by the meaning attached to funding in the relationships amongst participants. It was the aim of securing the next instalment of funding that kept the planning of evaluation going and that aim revealed how, in certain circumstances, power seemed to be manifested through the centrality of meaning attached to funding. Whilst this perceived meaning appeared very prominent, this finding has not previously been described when Atjonen (2015) assumed that power was mainly exerted through people. Thus, my study brings a fresh insight into the topic of PR in a DE context, expanding the understanding of evaluation in an imbalanced-power context. It also extends Westley *et al.* (2006) by adding a dynamic element of change on how the balance of power operates and it goes beyond the configuration of power

through people (Atjonen, 2015), emphasising the element of meaning through funding.

This study contributes to the literature on EP, by raising awareness about how PR are created, so that a practice-based evaluation keeps capturing how power and its relationships change over time; how tension is released, and how conflict is managed (Saunders, 2000). I argue that through a practice-based evaluation, the perception of the powerful funder is likely to be as a helper or partner, rather than a 'necessary evil'. Also, in a fast-changing society the individuals that currently hold the power will soon see it fading away. From a constructivist view, power is inevitable where two or more individuals co-exist, because there is a spontaneous force of willpower from one to another. This willpower is not restricted to evaluation settings, it occurs in all settings; in which humans interact.

A negotiation and co-construction of the perception of power seems paramount, however it is only possible if both sides of the relationship are aware, accept and capture how these are changing over time, and which circumstances are influencing that change (Shove *et al.*, 2012). My study extends present understanding of the debate on PR in evaluation, by arguing that in a funder-recipient context, power inequalities are predictable, and much more if the evaluative paradigm remains performance-based. However, as they are predictable they can be prepared and reflected upon, thus this thesis argues for and reinforces the call to consider a practice-based evaluation a meaningful and adequate approach to evaluate social interventions in a NPO domain.

7.3 Configurations of power relationships across the evaluation timeline

Power relationships (PR) came up predominantly in the analysis of participants' relationships with different people and were captured in three PR episodes, from which a specific configuration of power was generated. These configurations and the corresponding episodes are illustrated in Figure 7.3.

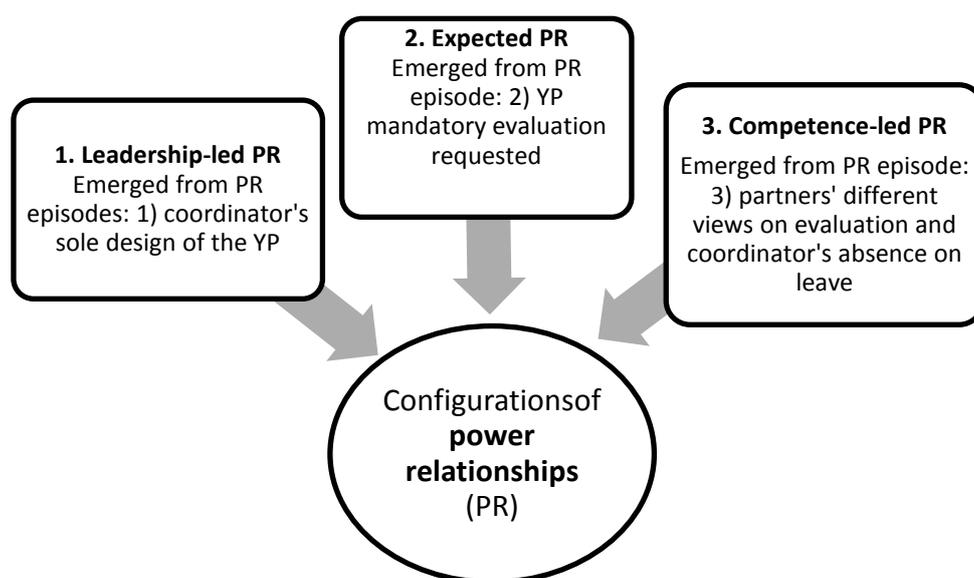


Figure 7.3: Configurations of power relationships across YPET

7.3.1 Leadership-led power relationship

The episode in which a leadership-led power relationship (LLPR) was observed was labelled as 'practitioners in the project design', to emphasise here the coordinator's sole design of the project (see Table 6.1; PCEs discussed). In recalling that stage of project design, particularly, the process of searching for adequate funding, practitioners referred to the coordinator's decision to design the project with DfID's funding scheme in mind (see section 6.4.1; disengagement). This seemed a challenging task, because previous youth work had to be adapted to fit into that scheme, so the question was: how to fit the

things that didn't fit and, more precisely, how to fit the evaluative element in this new bid? The perceived answer was Glenn's unilateral desire to implement DE with a focus on youth work activities, as illustrated:

We [NDEC] have developed those amazing simulations and we wanted to do more of them, so we wanted funding to do this. (Glenn, 1st interview, **emphasis added)**

Although Glenn, mentioned "*we*", practitioners reported barriers to participation (discussed in section 6.3) and sensed they had a lack of participation due to decisions taken in isolation. It appeared that, equally to Rip (2006), there was no general rule about whether some individuals contributed more than others, but some explanations arose, from Dale's description:

from my point of view, it would had been better for Glenn, while making the initial funding bid, to have a better communication with Jesse and me about how are we going to achieve the outcomes that **we** laid down for ...because **they were self**, we actually said that we will achieve those goals; **we were the ones that had put it down on paper.** (Dale, 1st interview, **emphasis added**)

The articulations of practitioners taken together suggested that the influence of Glenn's leadership was visible during the design of the project and that influence (manifested by their unilateral desire of project design) was interpreted as a power relation from one coordinator, as leader, upon practitioners. As this relationship appeared to be framed by the authority and hierarchy of the leader, I frame it as a leadership-led power relationship (LLPR), meaning that changes in practitioners' EP had not only been propelled by the funder's revised goals; but by their coordinator's practice of evaluation, which was perceived as "*self*" (Dale). Besides, practitioners had to readjust their EP, as Jesse noted:

If evaluation is 100%, **only 5% of effort goes to the evaluative bit**, the rest of the **energy** goes into the design and the running of the activities... **the evaluative side is [an] add-on**. (Jesse, 1st interview, **emphasis added**)

This re-balance of “*energy*” between the components of the project cycle, may be linked with the coordinator’s perception of the process of evaluation, as Glenn pointed out:

The problem with evaluation is when the tension comes at various times, because I had to get involved as manager. (Glenn, 1st interview, **emphasis added**)

Arguably, the coordinator’s attitudes and perceptions seemed to get inscribed in practitioners’ consciousness – influencing how they perceived evaluation. The PCE described as ‘practitioners’ participation in the YP design appeared to stay imprinted, thus shaping their perception and practice of evaluation (Shove *et al.*, 2012). Practitioners perceived evaluation as “*difficult*” (Dale, section 6.4.4), and this perception seemed to have consolidated over time, possibly because of their lack of participation in the project design - since they could not have engaged in the evaluation as they had not been involved in its creation. As such, a LLPR may influence how evaluation is perceived and practised amongst this group of practitioners. The power of a leader in charge of a DE organisation illustrates how a LLPR influenced practitioners’ perception and practice of evaluation across the YPET (Figure 7.1; PR-YPET).

My study observed an ‘internal’ layer of power emerging from the relationship between one coordinator, in the capacity of leader, and practitioners, as a team, illustrated, for example, by Dale’s account on the formulation of project goals solely by their leader. Contrary to Gong and Wright’s work (2007), mainly

focused on examples of PR with funders, this thesis shed light on an internal relationship where power was not only funding-related but also, leadership-led. The realisation of this internal layer of power, within the NDEC was a surprising finding, as previous literature assumed that power is exerted by the funders upon organisations (Carman and Fredericks, 2008). This insight challenged the view that most PR happened between two opposite layers of power – funder, as the powerful one, versus recipient, as powerless other. This thesis also challenges Alaimo's (2008) notion of a non-profit leader, whose features are summarised in Table 7.1 (overleaf).

Features of non-profit leadership (Alaimo, 2008)	How were these seen in a non-profit DE leader?
<p>Hierarchical position typically the highest-paid staff in a non-profit organisation</p>	Non-linear hierarchy not reflected in pay income
<p>Leaders driving the process leader plays an important role in shaping the organisation’s vision</p>	Funders driving the process, but this changes over time; whoever is in charge plays that role (e.g. practitioners or external evaluator)
<p>Managing the political environment challenge of being accountable to multiple stakeholders - some of them with competing demands for the leader and their organisation – to which leader balanced these responses while managing the strategic direction for the future</p>	Management of the political (and internal) environment is out-sourced: competing accountabilities and strategic decisions are supported by a strategic advisor.
<p>Incorporating organisation learning by influencing staff to achieve a common goal remain competitive, and reach their full potential</p>	Leaders influenced staff regarding their perception of evaluation; on some occasions by delaying their participation in the project cycle (chapter 6), which appeared to jeopardised the reach of their full potential.
<p>Decision-making and change Leaders must be able to handle the personal and professional discomfort that typically comes with the reflection and introspection necessary for effective cultural change</p>	Absent space for regular reflection

Table 7.1: Features of non-profit leadership in a DE setting (adapted from Alaimo, 2008)

By contrasting the features of a typical non-profit leader with those of a DE leader, I observed that the established idea that NPOs' leaders occupy the top-hierarchical position may change when applied to a DE context because, although they may be at the top position in the organisation's career ladder, the meaning assigned to their title may well be different from Alaimo's (2008). What does it mean to hold the top-position within a DE context? DE leaders tend to work in a state of constant pressure, therefore they are extremely demanding of their team (in theory, in a lower-hierarchical position), as well as partner organisation. For example, Glenn, one leader said "*what I learnt is that I won't partner with the other DEC [development education centre] anymore*" (Glenn, 2nd interview) and this may emphasise that the highly pressured funding context requires responsible strategic decisions regarding future partnerships. Consequently, leaders tend to be less available to intentionally involve practitioners in the strategic decisions, project design or evaluation (e.g. see section 6.4.1; disengagement). A possible explanation for this contrast was the fact that, in a DE context, funders appeared to be extremely demanding of DE leaders and these tend to replicate funder's attitudes onto their practitioners, as illustrated by Glenn's need to comply with funders' regulations, making sure their practitioners are aligned with them. An illustration of this leader's attitude was observed when Jesse and Dale asked to modify the design of the final report, adding more interactive resources and their leader refused, possibly influenced by the prior refusal from the funders regarding a film evaluation, rather than a formal report (section 6.4.3; resistance). Another possibility was that Glenn, as a leader was at the forefront of communicating with the funder, so if they were not pleased for any reason, Glenn would be the one to be blamed. This inference corroborates Alaimo's (2008) point about the challenge of being accountable to multiple stakeholders, some of them with competing demands for the leader and their organisation. My study suggests that within a funder-recipient based context, the way evaluation results are presented tend to persist, rather than to

change. It seemed that in the situation of producing a more informal and interactive evaluation report, the more practitioners requested a change, the more the current way of doing it (performance-based) seemed to persist and consolidate over time (Shove *et al.*, 2012).

Unlike Alaimo (2008) this study indicated that, in a DE context, funders tend to drive the process of management and leadership, with possible effects of shrinking leaders' competence and perpetuating a sense of meaningless evaluation onto their teams (Fullan, 2011). It is interesting to note, however, how change occurs in this process because in such an unstable employment setting as DE, whoever is in charge in a given moment plays the leader's role. As an illustration, my findings demonstrated that the PCE of the absent coordinator reinforced the role of practitioners in a novel way; because it was not the leader exerting the internal organisational control, but their team instead (See Table 6.1; PCEs discussed). This novel role of practitioners as leaders was visible during the planning stage (the leader was absent, so they acted as a leader); and unexpectedly during the implementation stage, when the evaluation was collaboratively-led by practitioners and external evaluator (e.g. for anonymity purposes, an example of this feature and episode was only available for examination).

Regarding Alaimo's (2008) feature of the management of the political environment, this thesis found that, in a DE setting, the internal environment plays a role as important as the political, because it is in the former that a DE leader can have control over their team, partnership and external decisions. Another contrasting feature was seen regarding the aspect of organisational learning, which in a DE setting was observed to be influential to whether practitioners reach their potential (or not) in their EP (further explored in section 7.4). Lastly, the feature around decision-making and change needs careful interpretation, as it is not quite

comparable in a DE setting, where a regular reflective space was not observed. In sum, this thesis argues that the concept of leader in a DE environment is bound to change over time, according to the surrounding circumstances.

7.3.2 Expected-led power relationship

The experience in which an expected-led PR (ELPR) was observed was the evaluation requested by funders. During the planning stage, participants knew what DfID was expecting in terms of evaluation, because they had been conducting evaluations for this funder's mini-grants scheme before, as Glenn, coordinator, acknowledged:

DfID requested a six monthly evaluation and project plan and then a year report. **But I would say that the evaluation of the mini-grants was quite superficial.** (Glenn, 1st interview, **emphasis added**)

Glenn reported that DfID's evaluation approach appeared flexible, and was quite relaxed about the experience with the mini-grants evaluations, because, possibly, the team knew how inauthentic that approach to mini-grants evaluation was, as Glenn and Sam, practitioners, recalled:

DFID, for me, is my perfect funder. (Glenn, 1st interview)

What I understand is that DfID used to administer their grants themselves and we [NDEC] were previously working with them and you send the financial report to DFID and someone will say, yes, fantastic or can you expand a bit on this section or check the figures and if there is anything wrong they ask you to resubmit. (Sam)

Sam emphasised that DfID's relaxed approach was because their grants were directly administrated, whereas the main change with the current evaluation was the appointment of an external GMA. In this respect, funders were asked to describe their relationship with NDEC, as Charlie, from the funder organisation, outlined next. As mentioned in chapter five,

I noted in my learning journal that the style of the funders' interviews appeared a bit manipulative and in Charlie's example they were specifically asked about the relationship with NDEC, and their answer seemed an abstracted reference to "*grant holders*".

I think our relationship with **grant holders** is slightly different from what [it] would be if it were with DFID in the **way we [GMA] act as intermediaries** between the project and the donor but also, from what I understand, **the requirements on us had changed very significantly, whereas in the previous administration there was not that many questions asked** in terms of the organisation's performance. (Charlie, **emphasis added**)

Addison, practitioner, outlined their view regarding this funder's evaluation approach:

Funders required evidence on the things we [NDEC] said that [we] will deliver, so we should expect that funder will then require for that evidence to be presented, with no surprise effect. (Addison)

Addison emphasised that the funders' rules were clear when the evaluation contract was signed, as well as, documentary evidence suggested that these were written down and made accessible to all practitioners and coordinators. Addison's account sheds light on the way practitioners assigned meaning to the funders' requirement of evaluation, as a supervisory, yet expected "*surveillance culture*" (Saunders *et al.*, 2011, p.4). In the same vein, another practitioner also noted:

If you look at the programs that **you want to get money for, you will go for funder's humps**; you can't tell them how to evaluate the project **they want you to produce**. (Dale, 1st interview, **emphasis added**)

For Addison, the funders' requirement of evaluation was almost a natural thing to do, while for Dale it was not credible, but observed as prescriptive and, maybe, intrinsically political (Cook and Leviton, 1991). Sam expressed another view:

People were quite **frustrated and nervous**, thinking if I am not getting it right, I might not get the funding. So, **there was a lot of pressure in doing it**

[evaluation] right and we had to do some workshops in evaluation techniques.
(Sam, **emphasis added**)

Whilst practitioners described experiences about their relationships with funders, I prompted them to express their views about what they thought was the main reason for the evaluation.

Jesse described that:

the **main reason for conduct[ing an] evaluation**, was not because there was a sense in the organisation that it was a good thing to do, **it was a requirement to do**. (Jesse, 1st interview, **emphasis added**)

The notion that PR with funders shaped the evaluation approach was described by Dale, practitioner:

So I assume that are some sort[s] of systemic disparity between the evaluation and the project funding. **If you are relying on the project funding, you will go by funders' rules**. (Dale, 1st interview, **emphasis added**)

Dale was explicit about the PR at stake when one relies on project funding and that illustrates the configuration of an expected power (an expected-led power relationship (ELPR)). Taken together, the previous account reinforced that within a funder-recipient based context, there were periods when the PR were characterised by the clarity of goals and expectations - termed as a balanced-power PR and other periods when PR seemed more imbalanced, illustrated for example by GMA's stricter evaluation approach (section 7.2; the context of a funder-recipient based evaluation). Within this expected configuration, I argue that the expectations amongst funders and coordinators seemed balanced, which was manifested through the acceptance of their PR. For instance, in a funding-dependent context, the expectations of a powerful funder were visibly influencing coordinators' and practitioners' practice of evaluation (section 6.2; the context of the YP: background information). As a result, this study indicated that various misperceptions of evaluation performance (and practice) arose, possibly resulting in a struggle to respond to funders' mandatory evaluation.

Previous literature assumed that NPOs can understand, assimilate, and thus practise a performance-based evaluation requested by funders (Carman and Fredericks, 2008, 2010); however, this thesis has been unable to support that view, partly because the perception of PR with funders appeared to generate internal pressure and urgency, leading to confusion about these requests. This atmosphere was illustrated by Sam's view on the frustration and nervousness shared by practitioners, which may suggest that a funder's request of a performance evaluation was expected, received, but not necessarily understood (these feelings are also illustrated by their need to attend evaluation workshops to "*get it right*"). Consequently, the effect of PR with funders may have led to reluctance on the part of practitioners towards the evaluation. This finding contrasts with that of Carman and Fredericks (2008; 2010) about the perceived linear way that a funder's request for evaluation seems accepted and how it prompted (or not) practitioners to move forward towards practice. Despite the expectancy that an evaluation will be requested by funders and responded to by practitioners, it is interesting to note how this expected PR may change over time – to a LLPR (section 7.3.1) or a CLPR (explained next in section 7.3.3). In this respect, my study highlights how configurations of PR may change over time, and how structures of power and their dynamics did not seem static, but evolved according to various circumstances, experiences, and contexts.

7.3.3 Competence-led power relationship

Some instances occurred in which a competence-led power relationship (CLPR) was observed; however, the most transparent could identify some participants, so for anonymity purposes, it was described in an appendix, available only for examination purposes. Nonetheless, two other instances are outlined next. The first was partners' different views on how evaluation should be practised. In the course of the project implementation they

appeared to hold opposing views regarding the approach to evaluation; for example, one assumed practitioners should meet deadlines and the other seemed more relaxed, letting the team manage their responsibilities. This evidence was partly captured during informal conversations and partly from my observations.

In last week's informal conversation, one coordinator mentioned that their partner was not doing anything on the evaluation, other than managing the project. And they were worried about the deadlines approaching.

November 2012

Extract 7.2: Extract from my learning journal, November 2012

These opposing views were more visible when one coordinator was absent on leave (see Table 6.1; PCEs discussed); which allowed the other to manage the evaluation according to their preferred approach. The PR observed was between both coordinators, and because they were at the same hierarchical level, power was exerted through their professional competence (to adopt one approach, rather than another). I termed this relationship a CLPR, because what was at stake was their competence, in a sense of having the required skills to practise evaluation. An important consideration was that because one coordinator was absent, the other was able to manage the evaluation more independently, which then seemed to cause further tension and modification of its practice, after the return of the absent coordinator (section 6.4.2; forcing). From this planning stage onwards, opposing views had to be negotiated, some persisted, and others faded away over time (Shove *et al.*, 2012).

The second instance occurred within a PCE of a coordinator's absence, in which practitioners had to reorganise their working routines, as Sam described:

What I also did was to set up a project management file and said to Dale, you've got to keep this up. When our coordinator comes back this will show everything that we have done at every stage. (Sam)

The coordinator's absence appeared to have allowed the space and opportunity for practitioners' competence to emerge and flourish, as they were accountable for the evaluation during that period. Possible explanations may relate to the lack of practitioners' experience in dealing with funders, thus they were able to choose and prioritise (their) competence, rather than the funding. Also my observations of this second instance suggested that the element of competence was brought to attention, because practitioners could prioritise their skills, ways of doing, thinking, and wanting (Shove *et al*, 2012). As such, within a context of an absent coordinator, competence appeared as a powerful drive in practitioners' evaluative practice – it seemed competence-led – rather than funding or leadership-led. PR were reconfigured when one feature of the DE context changed - in this episode, the absent coordinator – and a similar situation may well be transferable to other NPOs working environments, where one member of staff is made redundant, the project funding ends, or the contextual circumstances suddenly change.

The CLPR, experienced in my study, was mainly visible during the planning, implementation and reporting stages of evaluation, which reinforced my argument that the elements of funding and competence competed across the timeline (section 7.6; the change observed in EP). This study contributes to expanding existing knowledge (Saunders *et al.*, 2011) by proposing a competence-led layer of power, opposing the expected funder-recipient based context. What was also noticeable was the influence of opposing views in the collective practice of evaluation, particularly within a partnership, which shed light on how PR influences EP, through participants' perception. This will be explored in the next section.

7.4 The influence of power relationships in the practice of evaluation: the role of perception

To have a sense about their relationships, in the first phase of interviews, I asked participants specific questions about their role in the evaluation, as well as their views about others' roles. This section reports how participants' practice of evaluation appeared influenced by their perception of the surrounding power relationships (PR), as visualised in Figure 7.4, below.



Figure 7.4: Power relationships shaping evaluation practice through perception

7.4.1 Funders and Coordinators

When describing the evaluation planning, both coordinators acknowledged the issue of power attached to the funder's role. Glenn was particularly explicit about it:

Honestly I cannot stand GMA, because they are so rigorous and the attention to detail **drives me crazy**, because I do not have time to give them every little piece of information so I cannot stand it. **I hate the relationship with GMA, because it is a total power relation, where they say what they want and we do what they want.** (Glenn, 1st interview, **emphasis added**)

Glenn's account suggested an emotional aspect of their PR with GMA, showing concern in disliking their approach to evaluation; whereas the other coordinator outlined:

People sometimes see the evaluation as something that [the] funder wants...so I have to do it for the funder, rather than what can we learnt from that process and how to strength[ened] our own practice. (Eli, **emphasis added**)

During the second phase of interviews, I probed funders about their perception of coordinators' view of evaluation, as Charlie, from the funder organisation, outlined:

Well we only know what they tell us...if they tell us rubbish we don't have anything to work with. It might be great on the ground, they might be delivering, but if they not a) measuring it and b) **if they are not telling us in an intelligent way what they have measured, we don't have means to know.** (Charlie, **emphasis added**)

These accounts in combination with my observations suggested that power was most manifested through relationships, individuals' expectations from others, and also through what has been imposed and done. In challenging the view that a NPO leader practises their duties not (only) for imposition of the external forces (Fullan, 2011), Glenn's view illuminates the notion of a total "*power relation*", dominated by the funder, in which decisions on participation, negotiation, and ownership of the evaluation seemed to be absent. This absence may potentially lead to practitioners' scepticism in EP (further discussed in section 7.5.3).

The combination of coordinators' views with other participants' evidence on PR, suggests that funders' approach to evaluation appeared to have influenced coordinators' practice of evaluation (through perception). In other words, their narratives about power allowed me to infer that the funder's attitude influenced how coordinators perceived (and practised) evaluation. Equally, Charlie's account on what they expect to know from coordinators in "*an intelligent way*" reinforced a possible misalignment between both expectations. Similarly, to Carman and Fredericks (2008), I observed a perception of evaluation as related to a "*resource drain and distraction*" (p.34), alongside the recognition that coordinators' evaluation efforts were mainly power-led by funders' requirements. Whilst my study supports the idea that PR are visible in organisations with a potential less established evaluation culture (Schein, 1996), it sheds light on the influence of PR in the perception of evaluation across various levels of participants, as visualised in Figure 7.4, above. This section focused on how funders influenced the coordinators' perception of evaluation, which ultimately

shaped the latter's practice of evaluation. This dynamic persists and impacts the following relationship between coordinators and practitioners, as described next.

7.4.2 Coordinators and Practitioners

During the evaluation planning, coordinators and practitioners shared the view that the YP evaluation was being planned because it is mandatory, which meant that from a coordinator's perspective, they had to pressure practitioners to practise evaluation according to funders' requirements, as Glenn explained: "I drew an evaluation tool matrix and I listed every tool that we used and then I had to confirm that the team was using those tools".

In prompting practitioners about their views on being pressured, Jesse remarked that: "*we do have consequences if we have a very bad evaluation; it means that we don't have any more money*". So their illustration of pressure was manifested by a sense of having to comply with prescribed regulations, and at the same time please their coordinators, otherwise consequences may arise. The power through which a single funder required a mandatory evaluation implied an uneven relationship, which in turn, influenced the way coordinators perceived and practised evaluation. Moreover, coordinators' perception of a rigid evaluation was communicated to practitioners as the main reason to evaluate and that may have influenced practitioners' perception of their EP, as Sam, noted:

I think we are not quite good on evaluation, because Glenn was not as keen on that... honestly...**Glenn was more into the next thing, the next thing, the next thing.** (Sam, **emphasis added**)

Sam's view illustrated that one coordinator's attitude can influence practitioners' perception of evaluation, particularly when the latter had to respond to the powerful funder's request. It appeared that the current EP was being imposed by the funder, and sometimes reinforced by the coordinator, according to their convenience. A possible reason for this notion of

imposition may draw on the fact that, in a DE setting, evaluation tends not to be the core of practitioners' practice (Saunders, 2000).

Moreover, the coordinators' depictions indicated that past and current experiences tend to shape their perception of evaluation, and are likely to influence their practice. In turn, practitioners appeared to be influenced by the coordinator's amalgam of experiences, which suggested that the way one perceives evaluation is critical, as well as complex, and not only related to resources or time allocated (Saunders *et al.*, 2011). My study suggested that most of the time it was not what coordinators said that influenced the team, but what they did; as illustrated by Glenn's pressure on practitioners' use of evaluation tools, alongside the fact that Glenn "*was not that keen on evaluation*". Emerson's (1971 [1875]) motto: "*what you do speaks so loud that I cannot listen to what you say*"¹⁴ comes to mind, to illustrate that the practitioners' practice of evaluation seemed to have been influenced by what their coordinator has done, rather than what they have 'preached'.

Another illustration was observed when one coordinator acknowledged that they were detached from evaluation and did not know what was going on (discussed in section 6.2.1). Both examples ended up with the practitioners persisting and coping with evaluation, from which I infer that previous coordinators' practice influenced current practitioners' perception and practice of evaluation (the element of practice over discourse is detailed in 4.2; research using SPT). From a DSP view, the way perception and practice of evaluation intersects at a coordination level, influences practitioners' EP, as discussed next.

¹⁴ In 1960 President John F. Kennedy spoke at that Mormon Tabernacle in Salt Lake City, Utah and used this quotation that he attributed to Ralph Waldo Emerson (1971 [1875]).

7.4.3 Practitioners amongst themselves

When discussing the evaluation, the PCE described as the YP partnership in practice, shed light on the influence of PR amongst practitioners. From the combination of this episode with my observations, this study suggested that although evaluation is assumed as a social relation (Abma and Widdershoven, 2008) the issue of how PR influence and alter that practice in a DE context was left untouched. This literature assumed that an evaluation that is practiced in a responsive way “*help to unravel ambiguity and to enhance the mutual understanding between different stakeholders*” (ibid. p. 221). In contrast, this thesis shows that not only a funder and coordinator, in the capacity of leader, but also practitioners, appeared to not perceive this relational nature at ease, as Jesse illustrated: “*Once the evaluation is perceived as a **mechanical activity**, it actually loses what is all about and somehow that **feels unsatisfactory***” (*emphasis added*). In some instances, a sense of entanglement, increased ambiguity and misunderstanding were observed (illustrated in 7.3.1 and 7.3.2; LLPR and ELPR).

In sum, this section outlined how the influence of power travels across various levels of relationships in a DE setting, as visualised in Figure 7.4 above; power relationships shaping evaluation practice through perception. The main argument is that due to the particularities of this setting, practitioners found it difficult to understand the broader picture of evaluation, particularly, the theory behind different approaches and the various purposes of evaluation. As a result, they are likely to misperceive core issues about it, leaving an empty space for someone else’s issues to land (e.g. funders’ and coordinators’ ideas shaping their practice of evaluation). A practice-based evaluation would encourage the reflection needed to unpack some misperceived issues faced by practitioners on a daily basis. The discussion presented so far illustrates that power is a central aspect across the practice of evaluation and that PR tend

to shape participants' perception of evaluation. Therefore, PR are influential to the way evaluation is practised, and may generate some consequences for EP, as outlined in the next section.

7.5 The consequences of power relationships for evaluation in a Development Education context

Figure 7.5 visualises three consequences of PR for evaluation, the corresponding trigger episode observed, and how each consequence either shaped or changed EP. A note of attention is needed to alert that the figure below is stylised, as sometimes these connections were not as linear as may be suggested, and for that I added some illustrative messiness through the dynamic arrows. Each colour corresponds to a stage in the evaluation timeline (see Figure 7.1; PR-YPET).

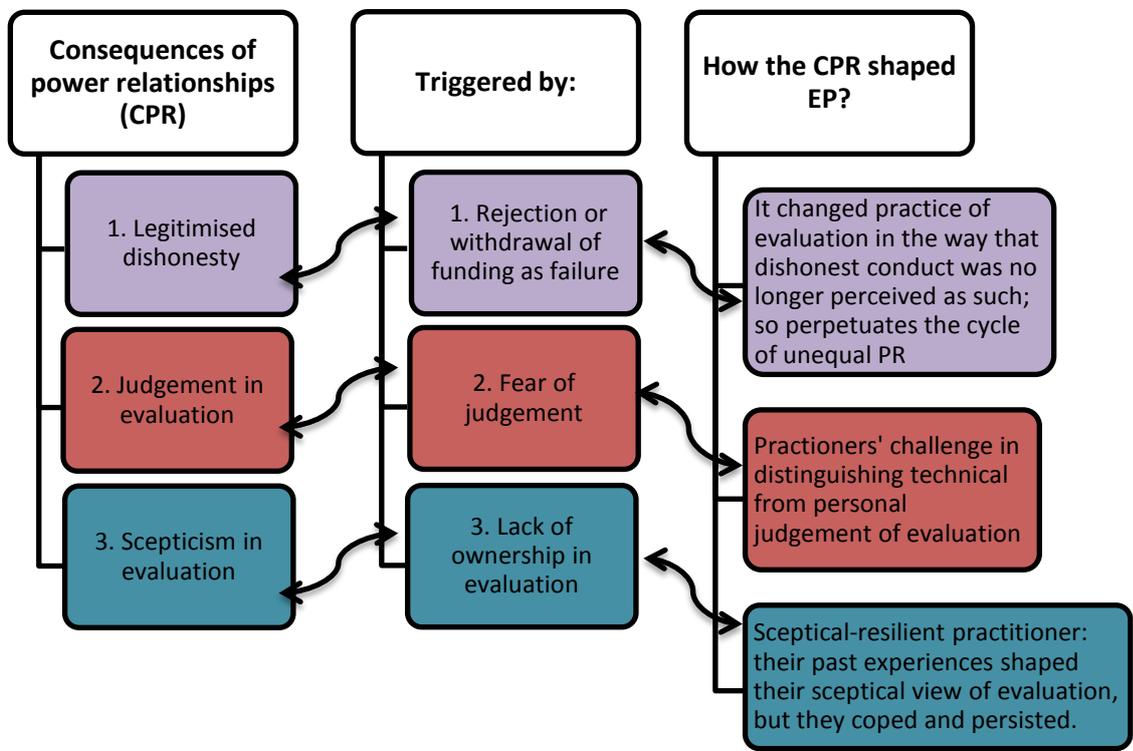


Figure 7.5: Consequences of power relations shaping evaluation practice in a DE setting

7.5.1 Legitimised dishonesty

In conducting an in-depth exploration of participants' experiences of evaluation, some argued that evaluation has to be conducted, because otherwise there were real consequences, as noted by Glenn:

If we don't do what funders want then we won't get the money. (Glenn, 1st interview, **emphasis added**)

Others referred to the possible consequences of their practice of evaluation:

By **poor evaluation**, I mean the demonstration that you were not as effective as you could have been **and that might have a negative impact on future funding.** (Addison, **emphasis added**)

If there were **no negative consequences to the evaluative process**, people will be **much more open to it, much more honest about it**, because they wouldn't feel that they will put the organisation down or all these **potential negative consequences.** (Jesse, 1st interview, **emphasis added**)

When probed about possible consequences of the evaluation, Charlie, from the funder organisation explained that:

Quite often there is a tension between DEC's [*Development Education Centers*] having to apply for funding to deliver some projects and what they want to do as an organisation. **They want to have the money to keep going** and in some cases I think they have to import their ideas into DAF [*Development Awareness Fund*] criteria and sometimes they cannot do it. That is one of the reasons for initial rejection when projects' objectives are not related with DAF objectives. (Charlie, **emphasis added**)

Participants' accounts opened up a sort of 'black box' in the exploration of issues of power within EP – its roots are so deeply ground within a funder-recipient based context, and relationships, that a poor evaluation with the consequence of a potential rejection or withdrawal of funding appeared to mean failure. This threatening scenario seemed to work as a powerful motivator to conduct the evaluation, by all means, especially, when losing the

funding (either by rejection or withdrawal), was reported by most participants as the worst case scenario.

During the evaluation planning, the main consequence was the rejection or withdrawal of funding, particularly because DfID had revised their expectations such that a strict performance approach to evaluation was in place. As the implementation of evaluation advances, coordinators and practitioners shared their concerns that the consequence of not conducting the evaluation was still connected with the vivid threat of having the project funding withdrawn, which would represent a failure, as a DE organisation. Towards the reporting stage, one connection was made with the consequence of losing the funding, as Jesse, remarked:

If an honest evaluation happens, it will be a danger because we would not receive another set of income which will put us out of business...we would have to close down because then we would not have any more money.
(Jesse, 2nd interview, **emphasis added**)

Jesse suggested that an “*honest*” evaluation would jeopardise the receipt of funding, thus coordinators tended to mask and hide the project’s weaknesses, which seemed to perpetuate their powerless status in this context. It appeared that practitioners were supposed to be ‘dishonest’ for the sake of assuring the funding. This study found that within a funder-recipient based context, the threat of losing the project funding, through rejection or withdrawal, was shared as a potential consequence, and as a result, to avoid that threat, some dishonest behaviour seemed to become acceptable – and legitimised. For example, during our informal conversations, practitioners reported the difficulty of having to gather data to ‘honestly’ demonstrate evidence to different funders. In doing so, they faced the challenge of having to “*slice and dice*” (Hoole and Patterson, 2008, p.98) data gathered from a single

activity in order to produce two or more evaluation reports for distinct funders, providing that each had different goals to match.

Similarly, to King (2007) my study indicates that practitioners' response to funders' power is mainly prompted by the necessity to secure funding. This legitimised dishonesty may prevent and avoid the failure of rejection or withdrawal of funding. Participants' assumption that an "*honest evaluation*" (Jesse) would certainly lead to failure highlights how power structures can be overcome. The flaws of the funding game seemed to have been disclosed and the instinct for survival arose, which in turn, appeared to have legitimised a more dishonest conduct on certain occasions. However, caution must be applied in this interpretation as other subtleties may not have been captured.

Literature on EP emphasised the need for openness and honesty as essential dimensions of a practice-based evaluation (Saunders *et al.*, 2011), however, my study challenges the idea that everyone will be honest when working in a context in which the importance of funding seems paramount, either to secure people's jobs or for their sense of identity as a practitioner. My case study indicates that in a power-led context, such as DE, the morals of honesty and openness may be manipulated to the point of legitimising a dishonest practice of evaluation. For instance, some practitioners' examples implicitly suggested that in a performance based-context, where funders' rules are strictly prescribed, coordinators and practitioners are encouraged to be slightly dishonest, for the sake of preserving the funding, to the point that slicing and dicing data seemed acceptably honest. That insight highlighted how practitioners no longer saw their actions as dishonest, but as reasonably legitimate, which may be explained by the need for further reflection and dialogue, ensuring that they can detach, for a while, from the evaluation process and reflect on its practice (see section 9.4.4; implications

for other practitioners). It is possible that this observed sense of legitimised dishonesty could be identified and traced back, particularly in a practice-based context of evaluation.

In this line, my study corroborates that of Saunders *et al.* (2011), but reinforces the challenges faced by practitioners in a power-led, performance and funder-recipient based context to manage their conduct of honesty in EP. This thesis extends the literature on EP by emphasising how the threat of rejection and withdrawal of funding triggers a sense of a legitimised dishonesty in a DE evaluation context.

7.5.2 Judgement in evaluation

This section draws on instances of expected PR (section 7.3.2) and barriers to practitioners' participation, to discuss how illustrations of judgement appeared to have influenced their EP. Specifically, it recalled two PCEs - DfID's change of goals and funders' different evaluation approaches (see Table 6.1; PCEs discussed) - to reinforce the fear of being judged in evaluation as one of the perceived consequences of PR amongst participants.

At the beginning of the project, practitioners appeared fearful about the strict changes in evaluation, initiated by GMA, which seemed to have generated confusion and misunderstanding in practitioners' EP. Besides, the Coalition Government had changed the DE policy, and as a result, DfID had to withdraw some of the other development education centres (DECs) funding, already received. An illustration of how that episode has impacted NDEC's atmosphere was recalled by Sam, practitioner:

The withdrawal of funding felt very un-transparent, very unfair, and **it also put a lot of fear into the rest of us [pause]I don't think I will ever trust DfID again.** (Sam, **emphasis added**)

DfID could have just phased out for a few years or they could have said actually we were not doing another round, but we will continue everyone

who's currently got funding...**but they just want to slash to show they were doing something**, that actually had an impact on some organisations in the country, **who then closed and lost key members of staff, knowledge and the skills of fifteen years of experience. How short sighted is that** [*pause*] you know. (Sam, **emphasis added**)

In order to fully explore this apparent fearful environment in which the evaluation was starting to be planned, I invited practitioners to describe their experiences of this stage. For example, Jesse described:

The sense of evaluation is one of judgement, that you are being judged, judging something, so it is not essentially about what can we learn in an open, **honest**, non-guarded sort of way, it's more about, what we are going to do now, **we are going to judge** if this project has been successful or not. (Jesse, 1st interview, **emphasis added**)

Expressions of frustration, confusion and uncertainty emerged, as a result of practitioners' apparent fear of being judged, as Jesse went on:

I think generally within the organisation, evaluation is seen as judgement. You will be judged on this at the end of the day, make sure that you pass the test, almost that sort of the approach. (Jesse, 1st interview, **emphasis added**)

These accounts combined with my observations suggest a link between judgement and fear—the sense of judgement seemed triggered by the fear of being judged (see Figure 7.5; consequences of PR shaping EP in a DE setting); and even by the fear of seeing the funding withdrawn at some point. The idea that an evaluative experience could be “*terrifying*” (Sam) gained momentum again, as well as the perception of evaluation as “*difficult*” (Dale), which the longitudinal evidence collected in this study suggests has not changed over time, rather it has consolidated. Arguably, funders' changes seemed to have stimulated subsequent changes in practitioners' perception of fear, fairness, and trust in evaluation, as illustrated by Sam:

We were submitting these documents requested by GMA [*Grant Management Agency*] and then it goes six months before they got back to you. So **that was frustrating and difficult** because we couldn't get instant feedback on

questions [*pause*] you know. **So we need to carry on, regardless of their feedback and then they say that it was not right...**well, we say, we have already done it now. We have delivered it, so it would be helpful if GMA clarifies. (Sam, **emphasis added**)

Interestingly, articulations of positive experiences, in general, and of evaluation, in particular, rarely appeared across the dataset, except one coordinator's reference to their self-competence:

I am very proud of what NDEC has become and it has become that largely because of me. (Glenn, 1st interview)

Although this evidence illustrates a positive insight of an overall judgement, caution is needed as it appeared to be self-centred and not made by other participants; however, documentary evidence indicated that the organisation kept going partly due to previous funded projects made by Glenn. On the contrary, Jesse articulated a negative view of judgement in evaluation:

I think **evaluation is a very heavy, onerous and negatively charged process** [*pause*] **and should be really positively charged**, but it has always been negatively charged. (Jesse, 2nd interview, **emphasis added**)

This study found interdependency between the idea of fear - risking practising evaluation without competence (as perceived in a performance approach) - and judgement - being judged on their identity as a coordinator or practitioner, regardless of the project's achievement. The link between an "*onerous*" EP and its judgement is situated around the power of funders, which is consistent with Taut and Braun's (2003) conclusion - an expected power relationship indeed generates fear, alongside misunderstandings about judgement in evaluation (see section 2.5.3.3; consequences of power relations in evaluation).

The notion of evaluation, as a difficult experience, spanned throughout the project timeline, particularly, regarding the contrasting evaluation approaches between the funders. This

contrast may suggest a “*doubting*” (Taut and Brauns, 2003, p.254) context for practitioners and coordinators within the DE domain, because from their narratives it seemed that when they started to perceive themselves as competent in understanding the basics of evaluation (and who has to do what) the funder-recipient context had already changed. My study also connects the issue of judgement in evaluation with competence, by suggesting that judgement was seen as a possible consequence of evaluation within a funder-recipient context based - either practitioner’s competence to evaluate was judged to be acceptable (Smith and Brandon, 2008) according to the funder’s objectives; or the evaluation was reported to the funder, and they will ‘evaluate’ it, as well as their proponent’s competence - a sort of meta-evaluation (e.g. from Sam’s view, the feedback process of GMA regarding the evaluation seemed judgemental and “*unclear*”). Two possible outcomes might occur: the judgement of the evaluation was positive and the organisation secured their funding; or, it was negative and they had to give back the funding received.

Judgement seemed one critical feature of how PR shapes EP in a DE setting, so I closely observed its meaning. To discuss the symbolic meaning of judgement, this research found that not only coordinators and practitioners seemed extremely defensive about their evaluation (Saunders *et al.*, 2011), but they also feared a negative judgement from the funder (Taut and Brauns, 2003). Hence, judgement in evaluation appeared as a consequence of power expectedly exerted from funders. What is curious about this finding was that coordinators appeared fearful in relation to the judgement of the evaluation, however, the outcomes to be evaluated (and judged) had been established by themselves; as practitioners have illustrated elsewhere: “*funders required evidence on the things we said that we will deliver, so when that evidence is required, there is no surprise effect*”. (Addison), and “*we were the ones that had put it down on paper*” (Dale). Similarly, to Saunders *et al.* (2011), this finding suggests that in a situation in which participants own what they designed, they rather

fear being judged and take a defensive stance – perpetuating the inequalities of power in this evaluative context. However, my study adds to previous work, by suggesting that practitioners do fear being judged and that fear has effects on their practice of evaluation, namely, the difficulty to distinguish between a technical and personal judgement of evaluation.

Moreover, some interdependencies of the funder-recipient setting are critical for the current discussion of the symbolic meaning of judgement, because evaluation, as a social practice, means that the elements of material, competence, and meaning link together (Shove *et al.*, 2012). As an example, the issue of judgement revealed that the symbolic meaning of evaluation seemed fearful and funding-dependent, which may have shaped how participants perceived their competence (section 6.2.1; funding dependency versus competence) - this insight sheds light on a constant changing of the practice of evaluation. Furthermore, as new elements of meaning entered the frame of EP –for example, judgement - others such as, practitioners' confidence in their competence, faded away (Shove *et al.*, 2012; see chapter three). For a given time during the YPET, practitioners' confidence in their competence was essential, whereas in other moments competence was replaced for others such as, funding or judgement. This alteration in the relationships between the elements of EP may suggest that participants' interpretations were unstable and changed over time.

Whilst this finding reinforced the centrality of the connection between judgement and competence, it is consistent with that of Saunders *et al.* (2011), regarding the continuous change of EP over time. Although, literature has recognised that practitioners tend to be judged on their competence to practise evaluation, I argue for the impermanency of that connection, as it evolves and another connection emerges (Shove *et al.*, 2012). This connection – judgement and competence - seemed to have changed over time, as illustrated

during the reporting stage of evaluation when practitioners had to collaborate and deliver the evaluation, regardless of being “*forced*” to engage, or even “*enjoying*” the process (Jesse) (section 6.4.5; collaboration). These changes in connections over time, add to DSP literature, that assumes that only connections between defining elements of EP - reviewed in section 3.3 - “*have to be renewed over time, so that stabilisation and routinisation are not end points*” (Shove *et al.*, 2012, p.24). My study showed that other types of connections may occur within, and alongside, the mainstream connection of elements of EP, as illustrated above.

This thesis extends the literature that conceptualises judgement within EP (Taut and Brauns, 2003; Saunders *et al.*, 2011) in two aspects: first, by reinforcing the fear of judgement as a consequence of PR, which may alter the practice of evaluation; and second, by underlining the impermanency of the elements of EP, and how they change over time, in the quest of accommodating contextual needs in each stage of the evaluation. Within that quest, practitioners’ sense of scepticism and lack of ownership in evaluation emerged.

7.5.3 Scepticism in evaluation: a sceptical-resilient practitioner

In describing the planning of the evaluation, I prompted participants about their learning from evaluation, inviting them to leave aside for a while the challenges already shared. As such, Tyler, practitioner, said “*I am not going into the philosophy of evaluation...I am only doing the work I have to do*”; whereas, their colleague, Jesse remarked:

I am not sure if I trust sufficiently well that the evaluation processes actually really reflect the learning, the journey, the process; so I am sceptical, I am still sceptical about that. (Jesse, 1st interview, **emphasis added**)

Whilst Tyler appeared unable to reflect upon evaluation, let alone their learning, Jesse emphasised the “*trust*” in evaluation. Tyler’s view may suggest that an instrumental view of

evaluation may increase in a context where a routine of reflection was lacking and where the practice of evaluation was neither valued nor reflected, as Jesse remarked:

If evaluation was not imposed by the funders, it will not be done. Or it will be done very superficially because it is not really valued. (Jesse, 2nd interview, **emphasis added**)

I also probed coordinators about their perception of practitioners' views of evaluation:

Some practitioners have not been giving the right attention or seriousness to evaluation. (Eli)

Also, there was a growing realisation that the evaluation planning needed improvement, consensus, "*attention*" (Eli), and "*to be valued*" – that there needed to be an overall, "*trust in the process*" (Jesse).

On the funders' side, the grant management agency (GMA) was demanding documents such as an M&E framework, but practitioners were unsure about what was required, as Sam and Dale, illustrated:

They do ask for a project framework when we submit the application but it is a very broad one. **It would be useful to find someone who knows this stuff as [a] specialist, who is being paid to do that, not an administrator, but someone who is willing to get it absolutely perfect** before you start delivering anything. (Sam, **emphasis added**)

When there is a bid, we have to know how many learning outcomes are going to be there, what exactly they [*beneficiaries*] will learn, how you're going to assess them before and after the process. That is what is needed for a funding application. But, **that can undo the creativity and can also refrain the thing that you are trying to achieve...because you are being prescribed to meet the targets**, so the assessment has to be **engaging enough for people to take part, but it has to be enough for the funders' assessment too.** (Dale, 2nd interview, **emphasis added**)

These views illustrate how GMA's performance-based approach to evaluation may generate a lack of ownership amongst practitioners about their EP. All of a sudden, everything appeared

to be questioned, their sense of self-confidence was low and they wished they had someone “*who is willing to get it absolutely perfect*” (Sam). As a result, practitioners asked funder’s advice on the documents for the project management, as Charlie, from the funder organisation, explained:

the project framework is developed once you have funding to do it and then alongside that there is a need to complete a M&E framework which states the objectives and indicators in the project framework and then add columns about your baseline, your source of data, how do you get to that process. Those were called **the reception document plans and are to be developed in the first few months.** (Charlie, **emphasis added**)

Whilst practitioners were having a hard time in understanding funders’ language and the specifics of each document requested, a sense of urgency was in the air, as the first few months had already passed. I noted on my learning journal:

Charlie seemed to speak another language (or a new dialect within the same language). Why is it that the participants’ expectations seemed misaligned?

December 2012

Extract 7.1: Extract from my learning journal, December 2012

Also, when I probed Charlie to further explore the evaluation, the straight reply was: “*I don’t know much about how it [YP] was as a whole; our interest is really to develop our funding schemes*” (see section 7.3.2 on the manipulative tone of funders’ interviews).

Later in the evaluation reporting stage, concerns such as whether the evaluation should be incorporated within the project completion report and whether videos were acceptable as acceptable evidence, emerged from practitioners’ narratives; and were manifested through contradictory opinions on how formal the evaluation report should be. Also, that ambiguity seemed to cause confusion amongst practitioners as they wanted to have a ‘say’ about the evaluation and for that reason had suggested a film evaluation as a more informal way of

reporting, but there was the issue of funders' acceptance of it, as discussed in section 6.4.3; resistance. The context of divergent expectations on how to report evaluation appeared to have prompted practitioners' reflection about the worth of their evaluative work, as Dale, expressed:

There is **very little authority brought to us** in that kind of formal evaluation.
(Dale, 2st interview, **emphasis added**)

Dale's view reinforced the idea of not being the owner of their work, which suggested that they lacked a sense of ownership and authorship. When I prompted practitioners about what they meant by not owning the evaluation, my observations shed light on the fragile strategic context of the organisation, at that time – a lack of financial sustainability alongside some practitioners' redundancies – leading to a sense of anxiety, that seemed to increase in a funder-recipient context (Greene, 1988). As illustrated in this section, the sense of lack of ownership in evaluation explicitly triggered practitioners' scepticism, as Jesse overtly described:

I have been involved in a project evaluation, three years ago. In the meetings they were all saying, what a bad poor project and, yet in the report, I read that it has been successful and I knew that it has not been successful at all. Somehow, **someone did the evaluation and came out positive and everyone involved knew that it was not [pause] so it came back, my scepticism.** (Jesse, 2nd interview, **emphasis added**)

From the combination of practitioners' transcripts with my learning journal, it seemed that their perceived belief of not being the owner of their work may generate scepticism in their perception of evaluation, which opens a possible avenue for discussion. Furthermore, the constant combination of sources of evidence was particularly enlightening, as it enabled me to contrast participants' transcripts across the two phases of interviews, and in Jesse's case, it was interesting to notice how their past professional experiences seemed to have shaped their

current sceptical view of evaluation. Past experiences are key for the inequalities of access and participation, as well as essential for the next experience over time (Shove *et al.*, 2012).

The idea of evaluation as “*dependent on an outside agent*”(Fetterman, 1994, p.4) is at the heart of practitioners’ sceptical views, therefore PR may explain some changes in EP, in terms of new actors emerging. The more disconnected, sceptical, and unskilled the NDEC practitioners appeared; the more qualified, external positions tended to appear (as was the case in terms of GMA, external evaluator, and strategic advisor). This feature in practitioners’ narratives seemed deeply informed by the beliefs that they did not have the necessary competence to evaluate, therefore, they could not own what they had done - because it was not “*good enough to please*” the funders (Glenn). It appeared that scepticism, in this specific DE setting, seems associated with the transmission of values, attitudes and perceptions across various levels of PR, but mainly across funders.

Likewise, Greene (1988), previous evidence suggested that when practitioners had not been heard in evaluation decision-making processes, the sense of their lack of ownership tends to increase; however, this thesis builds on the idea that practitioners’ sense of a lack of ownership results from an imbalanced context of power where different relationships are at stake (section 7.2; the context of a funder-recipient based context), combined with a sceptical view of evaluation, due to previous experiences. Jesse’s account suggested an illustrative influence of their reported past experience of evaluation and how it may have shaped their current scepticism. As these PR had to be negotiated (and change) over time, they may have contributed to a sceptical view of evaluation, in this DE setting, which relates to the notion of a “*sceptic factor*” in practitioners’ attitudes towards evaluation (Drewello, 2001, cited in Taut and Brauns, 2003); see section 2.5.4; the role of practitioners in EP.

One effect that predominantly arose was the sceptical view of evaluation, generated by participants' previous experiences of evaluation, as Jesse and Sam reported with an explicit sense of suspicion, almost distrust. Another possibility regards the negative effect/ influence of the practitioners' sense of lack of ownership in their practice of evaluation, was that it seemed to last longer in their minds than a positive one.¹⁵ In this scenario, Jesse's practice of evaluation was likely to be driven by previous experiences, overall sceptical. Although scepticism was most visible in Jesse's case, I suggest that past experiences of evaluation are likely to have shaped other participants' practice of evaluation.

I term Jesse's illustration to strive and cling to the evaluation, a 'sceptical-resilient practitioner' (SRP). SRP are project workers who seemed to struggle to translate the rationale and purpose of evaluation, endorsed by funders, into their daily practice. Some SRP, as Jesse illustrated, had previous (negative) experiences with evaluation that may have shaped their current view. Consequently, they appeared reluctant when first engaging with evaluation; but soon realised that they had to get involved in the process and coped with its challenges (section 6.4.4; coping). Equally, they sensed they did not have any option but to deal with their perception of lack of ownership, ignore their sceptical inner voice, and work collaboratively in the evaluation. Although Jesse endeavoured to keep the evaluation moving forward (despite being sceptical), their life path closely influenced the life path of their individual practice (Shove *et al.*, 2012) – discussed in section 3.5.1. This study indicates that when required SRPs rise to the challenge.

¹⁵ The reflection about past experiences, either positive or negative, and how they get inscribed in mind has been present in my journey, as a learner and practitioner.

Unlike Donaldson *et al.* (2002), who mentioned scepticism in evaluation as a problematic condition seen in evaluators, this thesis sheds light on the appearance of scepticism amongst practitioners – SRPs – not in a negative way, but as a consequence of a complex web of PR that have diverse configurations across the evaluation timeline (section 7.3; configurations of PR). It also raises awareness on the challenge of practising evaluation from a sceptical perspective, and for that, a practice-based evaluation is likely to ease that challenge, due to a constant reflection required over time (Saunders, 2000; Saunders *et al.*, 2011).

7.6 The change observed in evaluation practice: modification /stabilisation and competition

In relation to the theme of PR, the change observed in EP was through two effects: modification/stabilisation and competition. The first, modification/stabilisation was visible in a funder-recipient based context in which PR operates, and was illustrated by a cyclical dynamic of modification from a balanced to an imbalanced power (see Figure 7.2; balanced and imbalanced power across YP). The effect of stabilisation was observed when what was perceived as an imbalanced context, appears to be modified towards a perception of a balanced one, as soon as participants accept it, and cope with the contextual circumstances at stake. A second effect was illustrated by a competing two-way dynamic between an expected and competence led-power, illustrated by the elements of funding and competence, respectively. The thesis suggested that configurations of PR change over time, for various reasons, some related to issues of leadership - leadership-led power relationship (LLPR), funders' expectations – expected-led power relationships (ELPR), and competence of participants in evaluation – competence-led power relationships (CLPR). This study found that the change occurred between ELPR and CLPR was through competition; i.e. the core element visible in an ELPR, (the funding element) competed over time with the core element

of a CLPR, competence. An illustration of this competition was described and discussed in section 7.3.2, when the preponderance of funding seemed to have camouflaged the role of practitioners' competence in EP. Similarly, another instance later illustrated the preponderance of the element of competence within both coordinators. As they acknowledged having different approaches to evaluation, the apparent competition occurred within their competences, rather than with the role of funding. An illustration was one coordinator's relaxed approach to evaluation in contrast with another coordinator's opposing view (further evidence was provided in a confidential appendix).

Change is inevitable and has to be acknowledged and celebrated. The potential of the timeline to analyse the trajectory of change in EP is, that by analysing the past, I can make sense of why its meaning has changed, and what type of competences were required in the present moment that were not necessary in the past and so forth (see Figure 7.1; PR-YPET).

7.7 Summary and conclusions to the chapter

This chapter reported and discussed the theme of PR in EP. Findings and discussion were spanned across five sections as summarised in Table 7.2 (overleaf).

Power relationships (PR)	Findings emerged	Location
Context of a funder-recipient-based evaluation	1) A balanced-power emerged when funders' request was expected and previously agreed.	Section 7.2
	2) An imbalanced-power occurred when the funders' rules changed without forewarning, due to the Government's austerity measures.	
	3) A shift was observed from a balanced-power to an imbalanced-power context (as a result from previous change and resulting in subsequent changes over time).	
Configurations of power relationships across the evaluation timeline	1) Leadership-led power: internal layer of power is enacted by coordinator, replicating the funders' conduct.	Section 7.3
	2) Expected power: a mandatory evaluation is requested by the funder.	
	3) Competence-led power: power is manifested through competence; rather than through leadership or funding.	
Influence of power relationships in the practice of evaluation: the role of perception	1) Funders and Coordinators: The expected power manifested by funders shaped how coordinators perceived evaluation.	Section 7.4
	2) Coordinators and Practitioners: The leadership-led power manifested by coordinators shaped how practitioners perceived evaluation.	
	3) Practitioners within the partnership: relationships seemed power-led by the competence of some participants within the partnership.	
Consequences of power relationships for evaluation practiced in a DE context	1) Legitimised dishonesty: The rejection or withdrawal of the project's funding seen as failure, so a dishonest conduct appeared legitimate, to avoid the failure.	Section 7.5
	2) Judgement in evaluation: the fear of being judged as a consequence of PR in a funder-based context. Participants extrapolated the judgement from the project to the self.	
	3) A sceptical-resilient practitioner sensed not to own their practice of evaluation.	
The change observed in evaluation practice: modification/stabilisation and competition	1) Modification/stabilisation: funder-recipient based context in which PR operate in a cyclical dynamic of modification from a balanced to an imbalanced power.	Section 7.6
	2) Competition: two-way dynamic between an expected and competence led-power, illustrated by the competition of funding and competence.	

Table 7.2: Summary of findings on power relationships (PR)

The chapter concluded that different PR shaped EP differently. Whilst this conclusion does not aim to lead directly to the alteration of power structures, it allows novel insights into exploring/understanding how PR are created, maintained or modified. It emphasises that change subtly occurred, not only in the overall practice of evaluation but, also within each PR involved in that practice.

In this chapter, I argued that PR are influential to the way evaluation is perceived and practised by DE practitioners in a funder-recipient based context, which is, by default, a context where power operates between two extremes, funder versus recipient-organisation. My argument suggests that power can be manifested beyond these extremes and alongside an intermediate level where other subtle, yet unequal, configurations of power are developed (LLPR and CLPR). These new configurations of PR shaped the perception of evaluation, generating consequences to the practice of evaluation. This study contributes to expanding existing knowledge by proposing a competence-led layer of power, alongside the leadership and expected funder-recipient based context.

The issue of PR has been previously researched, but this thesis contributes to examining not the power issue itself, but how PR is configured in a DE evaluation context, and the circumstances in which they change over time (e.g. balanced to an imbalanced-power context). This finding reinforces the influence of the role of context in two dimensions: an external dimension acknowledging the powerful influence of funders in shaping EP that emerged through a funder-recipient based context. An internal dimension underlines the influence of practitioners' past experiences on their current practice of evaluation, with a

particular emphasis on how they carry EP to other practitioners or organisations; as had been carried through by some other carriers in the past. These insights on PR are an important addition to other studies focused on the practice of evaluation within a funder-recipient based context (Greene, 1988; Saunders, 2000; Saunders *et al.*, 2011; Atjonen, 2015).

From the analysis of the particularities of the case of a NPO in the DE domain, this study suggests that evaluation tends to be practised from a resistant stance, heavily shaped by PR with funders that triggered consequences for the practice of evaluation (legitimised dishonesty, judgement, and scepticism). The chapter also advocates that, although a competence-led approach was observed, the focus of power persisted as funding-based, probably, because the role and competence of individuals had been overtaken by the traditional preponderance of the role of funding.

This chapter intersects with the overall argument of the thesis by reinforcing a call for a practice-based evaluation, in which the changes of PR could be reflected, acknowledged, and expected in order to be accommodated. Within this practice-based approach, the temptation of dishonesty (even legitimised) could be avoided, due to the existence of a reflective space for planning the evaluation (regardless of the consequence of feeling judged or fearful). An evaluation where *practice* is pivotal distinguishes its operational/technical judgement, led by funders, from the personal judgement of the practitioners' identity (unconsciously led by them). Equally, practitioners would tend not to personalise the judgement of an operational evaluation, because they would, possibly, perceive it as a 'bridging tool' (Saunders, 2000), contributing to their

learning journey. From this practice-based view of evaluation, I propose that a sceptical-resilient practitioner can be supported in their daily practice through a celebration of their unique talents, which may imply an acceptance of other forms of practising evaluation. A possibility of an alternative form to practise evaluation would be to set out in equal terms with the funder in a relaxed atmosphere and having an enjoyable drink; which I argued¹⁶, may challenge the inequalities of power relations and incentivise a safe environment to approach evaluation as a learning process.

¹⁶ I presented this idea in a seminar given at the Department of Educational Research, Lancaster University with the title *"What if evaluation takes place seated around the table with a glass of wine?"*, in 16th March 2016. <https://www.youtube.com/watch?v=kx9h80FsOSY&feature=youtu.be>

8 THE CO-OCCURRENCE OF EVALUATION PRACTICE WITH OTHER WORKING ACTIVITIES

8.1 Introduction to the chapter

The chapter presents and discusses the findings that emerged from the analysis of the issue of the co-occurrence of evaluation practice (COoEP) with other working activities (WAs), in a Development Education (DE) working environment. It starts by outlining what participants did when they were evaluating the youth project (YP), as well as their other WAs, observed in a DE context. Section 8.3 describes three practice-change episodes (PCEs) and discusses possible influences on the NDEC evaluation practice (EP). Section 8.4 examines and discusses the change observed in EP over time, and section 8.5 summarises the findings and conclusions of the chapter.

The chapter begins with a visualisation of the issue of COoEP with other WAs and its emergence across the YP evaluation timeline. The rest of this chapter unpacks Figure 8.1 (overleaf).

FINDING: Co-occurrence of evaluation practice with other working activities (COoEP with other WAs)

COoEP COoEP with other WAs observed in practice-change episodes

C-EP Change observed in evaluation practice: by overlapping and dominance

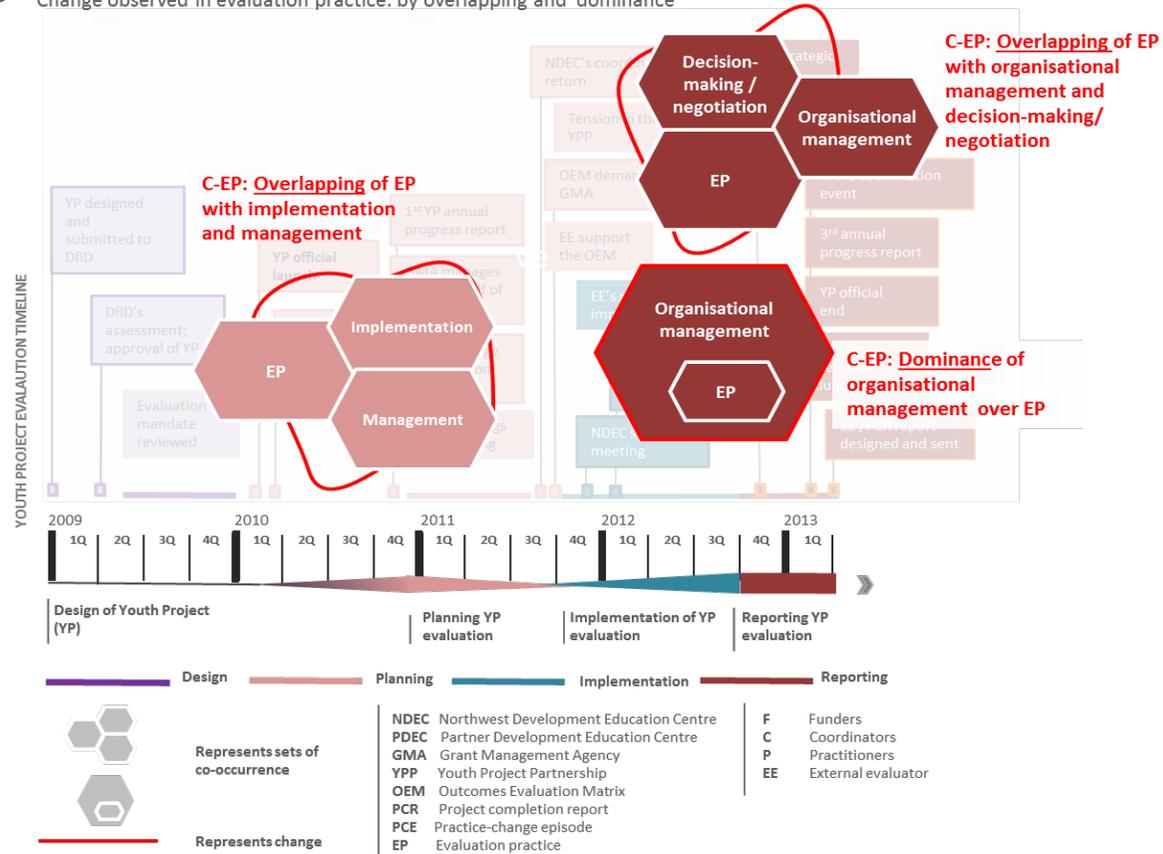


Figure 8.1: Co-occurrence of evaluation practice with other working activities in the youth project evaluation timeline (COoEP with other WAs-YPET)

8.2 Evaluation practice and working activities in a Development Education workplace

In describing their experiences of evaluation, two groups of participants mentioned the issue of various WAs occurring at the same time with their EP. However, for the purpose of this chapter, the term ‘participants’ refers to coordinators and practitioners (see section 5.3.4; composite profile of research participants) because funders WAs were not observed at this intersection. Also, I purposefully mentioned other WAs occurring alongside the EP to signal that this study primarily considers evaluation as a social practice; so other WAs, were not explored from this theoretical standpoint. For this reason, I use the longer term abbreviated to COoEP with other WAs, rather than simply ‘co-occurrence of practices’, as in the literature on dynamics of social practice (DSP) (Shove *et al.*, 2012). This section describes the dynamics of a DE workplace through participants’ portrayals of their working context. In describing their practice of evaluation, participants also mentioned other co-occurring WAs they had to conduct.

As chapter six and seven described the context and background of the evaluation in the first sections, this chapter starts with the factual evidence of the EP and WAs operating in NDEC’s working environment.

8.2.1 Evaluation practice

Participants were asked to fill in a form with some standard monitoring and evaluation (M&E) activities to capture their perception of the EP. This form was adapted from Carman (2007) and has been used in previous research in EP (see Appendix 5; evaluation form). As further explained in chapter five, funders had not been purposively invited to fill in this form,

because evaluation is already part of their legal duty (DfID, 2003). The summary of these evaluation forms is listed in Table 8.1, below:

	Evaluation and performance measurement activities	Coordinators	Practitioners
Typology of evaluation practice	Conduct formal program evaluations of projects/programs	2	2
	Use a performance measurement system	2	2
	Design program logic models	1	1

Table 8.1: Typology of evaluation practice (adapted from Carman, 2007)

Table 8.1 indicates that evaluation has been formally practised under a performance measurement system by both coordinators, and it is interesting to note that they reported contrasting views on evaluation, during the same interview (discussed in section 7.3.3; competence-led PR). For instance, one coordinator explained that they did “*not know much about evaluation...I should know, but I don’t know*” (Eli); whereas the other expressed concern about the financial sustainability of the organisation: “*the only thing helpful [pause] is really money*” (Glenn). Other concerns were also expressed:

I then withdrew from the project because I just had to focus on sustaining the organization [pause] to be fair I did write a couple of annual reports. **I thought it was Jesse that will do it, but they were so busy that I ended up writing a lot of it.** (Glenn, 1st interview, **emphasis added**)

The analysis of the evaluation form was combined with that of transcripts and observations to suggest that evaluation was mainly perceived as a performance-measurement practice, difficult to conduct and sometimes occurring at the same time as other WAs (e.g. writing of annual reports).

Regarding practitioners' views captured in the evaluation form, Table 8.1 has to be interpreted with caution, because only two practitioners, out of seven, referred to conducting 'formal program evaluations of projects', which highlights a potential ambiguity around the participants' understanding of the meaning of evaluation. I probed them about that ambiguity, asking for some examples from their practice, as Dale, shared:

I don't know exactly how evaluation influenced some of our procedures, but definitely I think we are more informed about what we want to, how we want it and the ways to go about it, as **things change very quickly and we need to adapt very quickly.** (Dale, 1st interview, **emphasis added**)

Dale's view seemed reinforced by my observations suggesting that participants had difficulty to directly enunciate examples of evaluative work, but in turn many times referred to change, particularly to the inevitability of instances of change, as well as to its unpredictability, as Sam illustrated:

I understand the political things that they were doing, as they want us to be supporting the developing countries and doing the direct aid. I understand that and **I understand why they have changed their goals.** (Sam, **emphasis added**)

Challenging times in NDEC seemed to have led to a sense of disorientation, poor financial planning, and uncertain future strategy, which had cut-across the EP (see Figure 8.1; COoEP with other WAs-YPET). From the analysis of the context, in previous sections 6.2 and 6.3, together with the current evidence reported, this study found that a multitude of concurrent WAs dwell in this specific DE context.

8.2.2 Working activities

Participants were prompted to report other WAs co-occurring with their EP, as listed in Table 8.2 (overleaf):

Typology of working activities	Working Activities (WAs)	Coordinators	Practitioners
Reporting activities	Produce reports for the board of directors	2	2
	Produce reports for funders about program activities	2	3
	Produce reports for funders about financial expenditures	2	2
	Produce annual reports	2	2
Regulatory activities	Conduct financial audits of your books	1	2
	Review program documentation (i.e., records, case notes)	2	4
	Acquire official licenses to operate programs	0	2
	Participate in accreditation processes	2	5
Monitoring activities	Experience site visits by funders or regulatory agencies	2	3
	Conduct performance reviews and evaluations of staff	1	3
	Conduct first-hand observations of program activities	2	3
	Monitor program implementation	2	3
Management activities	Assess whether you are meeting program goals, objectives	2	3
	Establish performance targets	2	3
	Engage in formal strategic planning processes	2	3
	Use a “balanced scorecard” management system	0	1

Table 8.2: Typology of working activities in a development education workplace (adapted from Carman, 2007)

Specifically, four categories of concurrent WAs were confirmed by participants: reporting, regulatory, monitoring, and management activities. Responses in Table 8.2 suggested that these four WAs have co-occurred, at some point with participants EP (see Table 8.1 above). A note of caution is that the analysis of Table 8.1 only captured responses from practitioners who perceived evaluation as a ‘formal program evaluation’, and in a broader sense, it may

have included other perceptions of what they consider to be ‘practising evaluation’. To overcome this issue, it was essential to closely contrast transcripts and observations for the analysis of this theme. Additionally, an interesting point to note when contrasting Tables 8.1 and 8.2. was that both coordinators appeared to conduct most of these WAs in co-occurrence with their EP, except for ‘the design of logic models’; which may relate to their reluctant engagement of practitioners in the project design (see section 6.4.1; disengagement) Regarding the reporting of WAs, participants mentioned they practised most of the activities, especially ‘reviewing program documentation’.

In terms of monitoring activities, participants reported that they conducted most activities on a daily basis; however, an interesting insight appeared regarding the ‘conduct performance reviews and evaluations of staff’ category, as these were expected to be conducted exclusively by coordinators, but practitioners reported to do it. This raises the issue about the reconfiguration of WAs within a DE context, especially when the human resources are finite, which may have explained the need for practitioners to conduct staff evaluations of their colleagues. Further informal conversations revealed that each practitioner had another practitioner allocated as their line manager, in order to ease the coordinators’ amount of concurrent activities. Another interesting finding was that although monitoring activities appeared as an established category, the research suggests that M&E are yet to be perceived as a systematic practice in this specific workplace (see section 6.4; typology of engagement of practitioners in evaluation), which may explain the difficulty expressed by Jesse:

Why are we now in the situation that the organisation might close? **Because real monitoring and evaluation have not been taking place, we have not really looked back as we went along.** Are we really monitoring what we are doing? (Jesse, 2nd interview, **emphasis added**)

Despite the coordinators' usual conduct of monitoring activities (Table 8.2), this evidence emphasised the unpredictability of change in a DE workplace, illustrated by the absence of one coordinator. As a result, Sam, practitioner, had to take on the responsibility (and change current workload) to accommodate the coordinator's activities and EP.

Another consideration was that the management activities showed that participants conducted most WAs, with the exception of the “balanced scorecard” management system'. Possible explanations for only one practitioner out of seven has reported this WA may relate to a possible ambiguity regarding the time-period covered in the form filling; participants might have thought about their past professional experiences, rather than the current EP (see section 5.4.1; technical issues encountered). The contrast in the triangulation of participants' WAs with their transcripts and observations appeared to suggest that, within a DE workplace, the assignment of daily WAs seemed led by urgency, availability, and physical presence, generating a situation of simultaneity, as Jesse, practitioner, explained:

I am not that motivated to focus [*pause*] it has been incredibly **difficult** because **lots of other things are happening at the same time**. (Jesse, 2nd interview, **emphasis added**)

This co-occurrence was also observed when the coordinator reported the need to write more bids “to keep NDEC going” (Glenn). A possible suggestion may be that management activities were concurrent with other WAs, alongside the evaluation (illustrated in a PCE discussed in 8.3.1). Also, the co-occurrence observed across management activities and others arose from the need to provide financial sustainability for NDEC in the long run, as well as making internal decisions regarding the current employment scheme. During an extraordinary meeting, further described in section 8.3.2, this WA was observed, particularly, regarding change to the organisation's legal status and conflicting views arose, as Dale, described:

There are **people managing the project, others implementing it and then conflict**. People are working towards **different agendas**, and at the end of the day they just want **some evidence of the project activities**. (Dale, 1st interview, **emphasis added**)

Dale's point was consistent with evidence in Table 8.2, regarding the 'review of program documentation', because this WA appeared as mostly conducted by practitioners, in co-occurrence with their EP. This co-occurrence was particularly noted by Jesse, during the reporting stage of the evaluation:

A decision was made by management, that the project needs to be **completed, delivered, reported, and evaluated** by the end of March. One could say that this is an indication, perhaps of how [*pause*] **that is an impossibility to do all of that, and to do it well. So that does indicate how serious or not the evaluation is**. (Jesse, 1st interview, **emphasis added**)

Jesse's concern sheds light on specific activities - negotiation and decision-making - that emerged exclusively from my non-participant observation of an important meeting; these WAs were visibly concurrent (see section 8.3.2; an extraordinary meeting). The findings presented in this section suggested the co-occurrence of aforementioned DE working activities across the YP evaluation practice.

8.3 The co-occurrence of practices through practice-change episodes

This section describes and discusses three PCEs, examining how the co-occurrence of WAs intersects and influences the EP across the project timeline, as visualised in Table 8.3 (overleaf).

Practice-change episode (PCEs)	Youth project evaluation timeline (YPET)	COoEP with other WAs observed
1. Different set of expectations in funder's evaluation approach	Planning evaluation	Evaluation practice, implementation and management
2. An extraordinary meeting	Reporting evaluation	Evaluation practice and organisational management
3.NDEC's change of office	Reporting evaluation	Evaluation practice, organisational management, and decision-making/negotiation

Table 8.3: Practice-change episodes in a Development Education working environment

8.3.1 Different set of expectations in funder's evaluation approach

The co-occurrence of WAs was observed when GMA was appointed to manage and evaluate DE projects. As GMA had requested evaluation plans to be completed and submitted from the beginning of the project implementation, practitioners had to provide those while implementing and managing the project. Also, that request was perceived as a change of expectations, because with DfID things had been rather flexible, as illustrated in section 6.2; context of YP. Various experiences were reported about GMA's role, for example, one coordinator, expected it to be aligned with the previous evaluation approach (by DfID), whereas practitioners pointed to the messiness in the organisation when the project started: *"we were preparing things and I would say that wasn't a great deal of planning for evaluation"* (Tyler). Tyler's point sheds light on a potential co-occurrence of the project implementation with its evaluation planning, which may have explained the lack of preparation mentioned. For instance, to implement the project, practitioners had to deliver activities, visit schools, and organise events, alongside planning the evaluation; which required a lot of internal organisation, as Dale pointed: *"we also had lots of discussions about management and organisational systems"*. As a result, coordinators realised that GMA's new

set of expectations would require a reorganisation of their EP, because they had to follow the Government's regulations to increase the focus on the planning – this was the main difference between them and the previous relaxed approach taken by DfID. From the participants' transcripts, it emerged that this increased planning and focus in evaluation seemed to have influenced their EP, as well as the conduct of other WAs, particularly the implementation and management of the project.

Regarding the management activities, practitioners monitored data; produced reports; captured images and videos to demonstrate results; managed the corresponding budget, and prepared M&E activities (section 8.2.2; working activities). Both of these WAs – implementation and management - were usually conducted by practitioners, sometimes collectively and at other times individually, but both co-occurred, alongside the EP, as Eli noted:

We used to be three persons to manage the organisation and now just me. So, **I am doing administration, finance, management, and cleaning**. Sometimes I feel that I have not been able to be in my role in some of the areas I have expertise [*pause*] **because I have been stretched too thin [*pause*] that probably is the main reason.** (Eli, emphasis added)

The experience of co-occurrence appeared to have increased Eli's sense of pressure to do a multitude of WAs, leaving behind other activities that they were competent, as it was Eli's evaluation expertise. From the perspective of managing the project, someone had to keep up-to-date with the funders and to check the project deadlines, because, at this stage of the implementation, practitioners had to organise an extensive data set to meet the deadlines agreed. As a result of the co-occurrence of the project management, implementation, and evaluation, some instances of pressure were observed, particularly amongst practitioners, as Dale outlined:

We have put a lot of time into things [*evaluation*] and we need to get a good output from it. **This pressure always forces you to** [*pause*] well, a long discussion about the right or wrong way to do it. (Dale, 1st interview, **emphasis added**)

Apart from these project-related daily activities, practitioners had other routinised tasks such as, the administrative work, the training of volunteers, the service in the resources centre, and other unplanned tasks that complicate the manageability of this landscape of co-occurrence. The context in which participants articulated the theme of COoEP with other WAs appeared dominated by tension and pressure, because while practitioners were implementing the project, they were also being asked to submit its evaluation plan (including an outcomes evaluation matrix). However, neither project partners were expecting to think about evaluation during the implementation stage, as Jesse noted:

we must all be accountable, but the way we may be accountable **changes the way which we do things**, sometimes therefore accountability becomes more important. (Jesse, 1st interview, **emphasis added**)

This PCE – of GMA’s different approach to EP - depicted that the perceived change in the funder’s evaluation approach was the inclusion of evaluation thinking across all stages of the project, rather than towards its end. Thus, to include that new expectation from the funder, various instances of co-occurrence between EP and other WAs may occur. The theme of co-occurrence of practices in evaluation has been previously acknowledged by scholars (Bonamy *et al.*, 2004; Saunders *et al.*, 2011), but scarce analysis on the dynamics of this co-occurrence, and its effects upon EP, have been brought to the debate. Whilst previous studies have recognised the co-occurrence of implementation, negotiation, and policy-making, possible implications for EP in the DE domain, have not been fully explored.

Additionally, current research on the trajectory of EP appeared to assume a certain linearity in the evaluative process (Bonamy *et al.*, 2004; Saunders, 2012). My study suggested that

COoEP with other WAs is a feature of a NPO, funded-recipient based context such as DE, where a perceived sense of urgency (e.g. illustrated by strict deadlines – Tyler and Jesse), tension (e.g. illustrated by participants’ opposing views - Dale), and pressure (e.g. illustrated by an unstable DE domain, with a few practitioners managing the multitude of WAs – Eli and Dale) arise. As a result, practitioners are expected to conduct every WA whenever it is needed, which is concurrently with their EP, as previously illustrated by Eli: “*I have not been able to be in my role in some of the areas I have expertise because [pause] I have been stretched too thin*”. This suggestion adds to the literature, by challenging the notion of a linear trajectory of EP.

Previous literature on EP also noted that evaluation tends not to be the practitioner’s exclusive practice, thus they will have the “*practicality ethic*” of individuals working in organisational situations (Saunders, 2000, p.14); however the translation of that ethic seemed to leave unanswered how other WAs were able to co-occur with EP. This co-occurrence is a critical feature of a dynamic routine of some working environments (Shove *et al.*, 2012), predominantly visible in funding-dependent contexts such as DE. My research indicated that when evaluation is practised in the context of COoEP with other WAs, it tends to alter, or even disappear, as Eli illustrated above.

Also, preceding literature implicitly assumed that practitioners need to participate in the evaluation process, to learn about and from evaluation (Saunders *et al.*, 2011). This assumption comes from the notion of community of practice (CoP) (Wenger, 1998), and from another critical assumption, that practitioners fully participate in evaluation, so that they can learn from it. My study challenged the view of evaluation as a CoP, by proposing that within a COoEP with other WAs, practitioners seemed to have another realm of preoccupations (with the issue of co-occurrence) to affirm their full participation in one activity or another.

Moreover, findings from my research underlined that although practitioners were practising evaluation, it was unlikely that they were fully participating in it, in terms of learning acquisition (Wenger, 1998) (section 2.5.3.1; participation of stakeholders). This challenge complicates things further when exploring practitioners' learning from evaluation. As argued in chapter six, their learning seemed fragile and inconsistent, because they faced barriers to participation, thus their engagement appeared non-linear. Possibly, practitioners in this setting have not learnt, as assumed in literature, but they might have gained awareness (Hall and Loucks, 1978) of how evaluation operates and changes in a context of co-occurrence. Furthermore, this specific working environment appeared to have altered the possibility of practitioners' full participation in evaluation, as well as in other WAs, jeopardising their learning from and in evaluation. In discussing what have they learn and how, this research found that although, the notion of learning, in literature is challenged, participants appeared to have learned how to work collaboratively in pressured situations and how to recognise difficulties and ask for support. In relation to how they have done so, although further research is needed to support this inference, the possibilities are that they have learned from evaluation by overcoming difficulties, making efforts to understand the process, while still maintaining some scepticism with regard to EP.

The current study expands that of Saunders and colleagues (2011), by suggesting that, a full participation in evaluation is at risk in a DE context where COoEP with other WAs operates. Possible effects were the difficulty of keeping EP at the core of a multitude of overlapping WAs (further discussed in section 8.4.1); the challenge of practitioners' full participation in EP, and finally, novel insights on learning acquisition (what and how). This discussion sheds light on how the funder's different set of expectations appeared to have influenced the change in EP. Whereas prior to the project, evaluation had been conducted in a more relaxed

approach by DfID; after the funder's revised expectations, evaluation was being practised under GMA's stricter approach.

8.3.2 An extraordinary meeting

During the reporting stage of the evaluation, there was a PCE in which the co-occurrence of management activities and EP was observed at: a NDEC's extraordinary meeting, with trustees, staff, one coordinator and facilitated by Brook in the capacity of strategic advisor (see 5.3.5; composite profile of participants). The main purpose of this meeting was to decide the strategy regarding the organisation's legal status, its financial sustainability and the employment schemes needed. The latter was observed to be particularly relevant in a period where some practitioners were being made redundant. During that meeting, I observed anxious practitioners and had a sense of the multitude of diverse activities they were asked to do, from cleaning the office to writing NDEC's statutory statement. I also captured in my field notes their interpretations of the strategic possibilities for the organisation, which seemed to carry veins of frustration:

Lost in transition

Today, practitioners told me about the DEC's lack of strategic direction and how demotivated they were. They mentioned that they will probably merge with another organisation, because the reality is that various centres are closing down, due to the Government's austerity measures. But they seemed suspicious about this option and I have not heard anyone else speaking about that. One practitioner also mentioned that they either merge or close down and they looked so frustrated that they even confessed to be job-seeking. How can they focus on writing the evaluation report? What strikes me the most is their urgent need of direction, which seemed the main reason for their anxiety.

NDEC

Extract 8.1: Extract from field notes, November 2012

Whilst I observed participants' frustration and anxiety, there was a sense of dominance (discussed in 8.4.2) from the management activities with repercussions on the EP, illustrated

for example, by the need to place evaluation aside for a while, due to NDEC's expectation that all staff would be involved in this period of critical decision-making. In this regard, practitioners were being invited to participate, to have their voices heard, to engage and fully participate in that strategic moment in the organisation's life path. As a consequence, their engagement and practice of other WAs, namely, EP, was dominated by management (and partly decision-making) activities as discussed next.

The findings that emerged from the theme of co-occurrence showed that within a DE working environment, various WAs co-occurred with EP, regarding participants' type of skills, priority of the activity and time and space allocated to practise it (Shove *et al.*, 2012). Regarding the type of skills, referred in the thesis as competence, Tables 8.1 and 8.2 indicates that within this specific context a co-occurrence of competences may also occur; for example when the competence based on performance measurement systems appeared prevalent. This insight contrasts with the literature promoting interdisciplinary competences in evaluation (Trowler *et al.*, 2012; Simons, 2015) by emphasising that in a context where various competences intersect, still the performance-based one triumphs. Although this possibility was partially observed across the evaluation, further research on these implications is required. In relation to the priority given to a specific activity, rather than another, my observations suggested that in times of financial scarcity (e.g. when funding is about to finish), priority is given 'to just do the next right thing' (Kelly, 2008), shrinking the coordinator's discernment and long-term vision. However, it appeared that whilst the coordinator and trustees, understandably wanted to prioritise the management of the strategic future of organisation, at least one practitioner seemed keen to keep practising the evaluation and their other daily routines. Some others appeared more concerned with their employment status as they were already job-seeking; therefore evaluation was at the end of their priorities list. Although, these findings highlighted a situation of apparent dominance of management

activities over the EP, it was not a direct effect of COoEP with other WAs, because a slight tone of competition emerged from informal conversations held with participants, which was captured in my learning journal:

Addison told me that the most important thing was to know what will happen next with NDEC, especially because they have other collaborations as they are self-employed, and have to account for that. Sam has already been made redundant and currently job seeking, so has no concerns with evaluation, whatsoever. Jesse seemed to be conducting an internal battle; on the one hand they want to show respect for the colleagues and show support in this critical period; on the other hand, Jesse was ultimately responsible to ensure that the evaluation deadlines were met. Also they were closely collaborating with Brook, in their capacity as external evaluator, so possibly Jesse wants to show Brook that they were on track with evaluation. What are the odds that Brook would be simultaneously facilitating this meeting?

November 2012

Extract 8.2: Extract from my learning journal, November 2012

Previous work on EP advocated that the meaning practitioners give to evaluation should be taken more carefully into account, because their understandings affect EP (Seppanen-Jarvela, 2004). My research suggests that the meaning practitioners give to evaluation may relate to financial, personal, and emotional aspects that tends to be kept hidden in the current research of large and politicised NPOs (Carman, 2007) (see Table 4.2; overview of the research gaps). This was illustrated by the type of concerns practitioners mentioned during my observation of the extraordinary meeting combined with previous analysis of their transcripts, particularly in relation to the priority given to funding since the project design (Jesse); the personal responsibility in providing financial security for their families (Addison), and the emotional sense of not being “*really valued*” in their profession (Sam). These aspects – financial, personal, and emotional – may have influenced their understanding of evaluation to the point of changing their practice (discussed in 8.4). Although my research supports that of Seppanen-Jarvela (2004), in relation to the need to take practitioners’ meaning of evaluation into account, financial, personal, and emotional aspects should be considered vital in the attempt to understand where the meaning of evaluation comes from.

In observing this extraordinary meeting as a PCE, the time allocated for practising evaluation, and the other concurrent WAs, emerged as critical, especially towards the end of the evaluation, as Dale, outlined:

I think just time in itself was a big stress factor. I have done some other activities which have been very [*pause*] you know, **lots of time and lots of effort.** (Dale, 2nd interview, **emphasis added**)

Other participants' accounts referred to the aspect of time, as illustrated by Glenn and Jesse, respectively, with my emphasis: "*we have not much time to evaluation*"; "*other things in the organisation are happening at the same time*". In relation to the issue of time, my findings suggested that participants seemed to not have other options rather than to share practices, responsibilities and other WAs. The process of how practices come to be shared has been explored in the literature of social practice (Turner, 2001); however the emphasis seemed to be on how past experiences shape current practices, and less on how a situation of co-occurrence results in practices or WAs being shared. Another interesting point debated in the same literature was whether a community, group or organisation is characterised by the practices in which their members engage (Turner, 2001). In chapter six, I argued that the process of an inconsistent level of participation in evaluation influenced participants' perception and their practice. Moreover, this study sheds light on their apparent ambiguous motivation in practising evaluation, and whether they only do it "*to please the funders*" (Addison), or because it has been "*dictated from above*" (Dale). The PCE discussed in this section thus far, highlighted that in a situation of co-occurrence, many other activities overlap, compete and dominate the practice of evaluation. This finding allowed me to infer that my participants may have recognised themselves as engaged members of a global and changing community, as DE, rather than a single evaluative community. The concept of community here means that connections through which practices develop and circulate, and

by means of which they engage practitioners, do not necessarily map onto organisational or institutional structures (Shove *et al.*, 2012). That was visible in the PCE of that extraordinary meeting, in which I observed participants trying to make sense of what was their closest ‘tribe or territory’ (Trowler *et al.*, 2012), within a context of a multitude of activities. My study contrasts with that of Turner (2001), by suggesting that in a situation of COoEP with other WAs, organisations in a DE context seemed not only characterised by the practices their members engage in, but rather by an amalgam of different, concurrent, and sometimes ambiguous practices or WAs, that their members may or may not engage with over time (see section 6.4; typology of engagement of participants in EP). My research also contrasts with the literature advocating that community and practice constitute each other (Wenger, 1998), by proposing that a community can be constituted by a combination of practices; my case suggested that, in a DE context, a global and changing community was configured through the COoEP and a multitude of other WAs. The relevance of the issue of time seems crucial within the discussion of the co-occurrence of practices, but still sheds light on how these perceived resources (competence, time, and space) appeared taken for granted in some studies on how NPOs practise evaluation. This thesis highlights that the notion of ‘not having time’ arises from the construction of evaluation as a power relationship in which funders dictate the ‘available time’ to an evaluation and organisations have to comply accordingly.

This discussion focused on the type of skills to practise evaluation and how different competences sometimes seemed to co-occur. It also mentioned the issue of priority given to a certain activity in a COoEP with other WAs and how that meaning appeared to have influenced and changed the EP. From my observation and informal conversations of an extraordinary meeting, most staff was focused on their professional situation and NDEC’s future strategy, leaving the evaluation aside, which generated a different – dominated - EP towards the end (discussed in section 8.4.2). Lastly, I discussed the issue of time, as a

resource allocated for EP, and how time was perceived within a COoEP with other WAs. The next sub-section describes and discusses the issue of space, and its possible implications for the EP.

8.3.3 NDEC's change of office

One of the outcomes from the extraordinary meeting was to change the office location, so the team moved to a shared building for NPOs, occupying two small rooms. The new space was smaller than the previous and that may have shaped how participants related and communicated amongst themselves, as well as how focused and productive they were. The move was observed to take place in March, 2013, which meant that they had to pack, store, sell, and transport all furniture during the previous weeks, resulting in the postponing of the final evaluation report. I captured one instance from this period below:

Another beginning

The move to the new office is in March, so I offered help with my car, but they said that they were already sorted for that day; so in turn we spent all the morning tidying up and sorting things to sell. Whilst we were doing it, we chatted about the new timing for delivering the evaluation and one practitioner was worried that it will all end up on their shoulders. They were also worried about the practicalities of the new office's space: computer, internet connection, and telephone landline; as those were the basics things they need to keep working. As they were busy, we agreed that I will interview them again next month, rather than in the next week.

NDEC

Extract 8.3 Extract from my learning journal, January 2013

As Extract 8.3 highlights, the organisational management activities co-occurred with the continuum of the evaluation, which may have influenced participants' behaviour; their communication between the two rooms in the new office; their negotiation of the working spaces' arrangements, and the provisions to receive visitors. That influence was also

illustrated in one incident, during the second phase of interviews, as the following extract illuminates:

Where are we sitting?

It was the 21st March 2013, at DEC's new office and I started to interview Glenn downstairs in the main foyer, because there was no private room for us. I recorded eight minutes of conversation when other people arrived and sat nearby; their noise was so disturbing that I stopped recording. We then decided to come back upstairs and Glenn searched for an empty room; a kitchen; a store; whatever may be, but there was no space for us, as those were all locked out. So I was invited to sit on the floor near the hallway stairs. I was obviously surprised but sat with Glenn, pushed the record bottom and went on with the interview. In the meantime, other colleagues were getting in and out from the office, not grasping what we were doing sitting on the floor, a rather odd place for a talk. At some point, these colleagues wanted to join us and I was embarrassed looking at Glenn, who told them that we were in the middle of a research interview. I recalled my embarrassment, as I wanted to keep going with the interview, so I had to remind them that we were having a private conversation; but how on earth I will do that if we were seating in a public space. I did not want to be rude, so I coughed and Glenn eventually told them that we had to keep going. I remembered their surprised faces, and my thought was about having fun with my research. I also thought about space and how an 'unfamiliar' space shapes ones' working practices. We were working on the floor, whereas for colleagues passing by, the meaning assigned was different - definitely not a working practice meaning.

NDEC (new office)

Extract 8.4 Extract from my learning journal, March 2013

The instance described in Extract 8.4 may illustrate that changing workplace (organisational management) can influence how the coordinator relates and assign meaning to their experience of evaluation. However, another earlier incident described in chapter five (see Extract 5.2), suggested that the coordinator (and possibly practitioners) were already relaxed during the research interviews, because in the situation described in Extract 5.2, Glenn interrupted the interview, left the room, went upstairs and picked up the phone. Similarly, in the description in Extract 8.4, we could have sat on the floor for various reasons, first of all, there was no other space, second, they were observed as having a relaxed attitude about the

interviews, and third, possibly they may have perceived evaluation as a non-priority practice. In contrasting Extract 8.4 with previous evidence, particularly with Extract 5.2, findings suggested that the priority given to evaluation did not seem high, possibly due to other WAs and contextual constraints (discussed in sections 6.2 and 7.2). As a PCE, the change of office location illustrated that participants had to cope with unpredictable situations; yet, in a situation of co-occurrence, some implications seemed to arise for EP, which are discussed next.

The combination of PCEs described in this section with evidence from chapters six and seven, allowed me to infer that the way evaluation has been practised, within a DE working environment, seemed mostly unreflective and lacking in awareness. Specific illustrations were Tyler's sense that there was no need to think about the philosophy of evaluation, and Jesse's point on the organisation's undervalued perception of evaluation. Possible reasons for the unreflective nature of EP relate to the pressure of the day-to-day work, which meant that practitioners were unable to stop, review, and reflect. Consequently, and because they appeared to work on automatic pilot, the working day did not seem long enough for them to reflect and to gain awareness (of how evaluation was being practised and changed over time). Another potential connection with earlier findings may be that, because of the lack of reflection, participants did not seem confident in their EP, and even expressed anxiety and fear when engaging with evaluation (see section 6.4.3; resistance).

Previous studies have recognised the key role of reflection in improving EP (Saunders, 2000; Saunders *et al.*, 2005, 2011; McCluskey, 2011); however, my study adds to that of Saunders (2000), by identifying that participants were not used to engaging in any type of reflection, due to a lack of time, resources and an increased demand of funders' requests. My research indicates that even in a COoEP with other WAs situation, where reflection was perceived as a

luxury of time, practitioners were still able to plan evaluation without their coordinator (absent on leave), to produce and deliver complex documents, such as outcome matrix, logic framework and annual progress reports to funders. Practitioners were also mindful of having to engage in evaluation, so they overcame an initial disengagement, during project design, and were perceived as forced, resisting, coping, then collaboratively engaging and working with the external evaluator towards the end of the evaluation.

The interesting insight about this association of findings is that participants (and here I mean the three groups) appeared unconscious of the overall continuum of their practice, reinforcing the notion that practices can co-exist and co-depend in ways that they hardly recognise (Shove *et al.*, 2012). More precisely, practitioners and coordinators seemed to have gained that awareness due to a retrospective reflection, facilitated by this research process. After the second interview, where we were seated in the foyer, Glenn confessed to feel empowered after the interview, because it was the first time for ages that they had put time aside to reflect. By suggesting that these DE individuals were initially unaware of reflection, as a tool to enhance evaluation, this study underlines how small NPOs tend to practise an unreflective evaluation, in their day-to-day life.

Saunders *et al.* (2011) suggest that reflection about practitioners' EP enhanced their awareness of the changes occurring in that practice over time. My study indicates that unintentional reflection of evaluation appeared to also enable participants to reflect back and forth across the project timeline, in order to see connections, they were not aware of before; as well as express experiences that challenge participants' views of evaluation. For instance, this was illustrated by Glenn's intention to become more reflective and "*honest*" about their practice (which was supported by the retrospective dimension of the research). This research is consistent with that of Saunders *et al.* (2011) in reinforcing the need for reflection, for an

increased awareness of change in EP. Besides, in the current study there was an unexpected finding of the benefits of reflection generated by the research process, which was the realisation by some of the research participants that a reflective practice resulted in an increased engagement in the evaluation towards the reporting stage. For example, the sense of this benefit emerged from an in-depth longitudinal account of one coordinator's experience of evaluation over time. The stage of reporting evaluation seemed an emotional one for this participant and the first moment in three years where thoughts about learning from the practice of evaluation emerged. Eli noted: "*what can we learn from that process [...] to strength [ened] our own practice*"; however, their reflective practice was limited and infrequent, which may jeopardise their learning over time. On the occasions where participants did reflect, they were able to articulate the benefits of reflection by recognising any learning, confidence and awareness of EP. When prompted by the research process, participants were able to reflect on the contextual constraints and challenges of their EP. Thus, I argue that, in a DE context, had reflection been present at any time evaluation was enacted, a COoEP with other WAs, would have been less likely – because participants would have allowed themselves time to reflect and accommodate the multitude of practices and WAs according to NDEC's resources allocation (e.g. the nature of change due to the austerity cuts; the assumption that all the work is manageable under the budgetary and staff skills circumstances).

This study contrasts the assumption that reflection is present at any time a group of practitioners enacts a practice (Saunders, 2000). The case under investigation showed that, in a DE context, the awareness of reflection, as a key aspect of EP, is yet to be internally conquered possibly because of the inconsistent level of participation of practitioners in evaluation (chapter six) or the novel configurations of power relationships (chapter seven). These particularities may apply to wider contexts within the non-profit sector.

This research revealed an unexpected finding in terms of the centrality of reflection and its benefits for the practice of evaluation. Past research on DE found that although reflection can act as an enabling tool to evaluation (Bracken and Bryan, 2010), no substantial reference of time, context, and frequency of reflection was made. My research sheds light on the role of a retrospective reflection in increasing practitioners' awareness of their trajectories of evaluation. Recognising its critical role allows for a refined notion of engagement, participation, and ownership in EP to come into play; supporting this thesis' argument for the need for a practice-based evaluation, in which practitioners can act as essential carriers to humanise and assign meaning to EP - a meaningful practice of evaluation. The challenge is that within a DE domain, where EP overlaps with other WAs the possibility of practitioners' recurrent reflection seems restricted. Not surprisingly in the current performance-based evaluation operated by funders, the role of reflection appears jeopardised, so a proactive stance from practitioners in small NPOs would support a practice based-approach to evaluation.

This discussion sheds light on how the change of office seemed to have influenced participants' predisposition to practise evaluation, possibly due to a limitation of resources and funding. In summary, the practice of evaluation within a COoEP with other WAs context appeared to have been influenced and changed after three PCEs episodes: first, when funders' expectations to evaluation were revised (section 8.3.1; GMA's revised evaluation approach), second, when participants' professional concerns co-occurred with NDEC's management activities (section 8.3.2; an extraordinary meeting), and third, when the working space was altered (section 8.3.3; NDEC's change of office).

This section highlighted how the aforementioned episodes appeared to have influenced change in practitioners' EP. Next, I examine how that change has occurred and possible implications for the EP.

8.4 The change observed in evaluation practice

As noted in chapter two, this study builds upon present understanding about how DE NPOs evaluate their social interventions, however previous studies have not examined the implications of the COoEP with other WAs. Therefore, the issue of change influenced by the COoEP with other WAs appeared under - researched in the evaluation literature, particularly in the DE domain (Table 4.2; overview of research questions). In contrasting some of the literature reviewed in chapter three, this section discusses instances of change to respond to RQ3: How has evaluation practice changed across the youth project evaluation timeline?

8.4.1 Overlapping

Overlapping is one form of selection and integration of elements of practice that can be dependent on each other. A DSP lens of evaluation considers that the way elements of practice and practices interweave can be through overlapping, modification or adaptation (see section 3.5.3; the effects of the co-occurrence). An illustration of a situation of overlapping was when GMA's evaluation approach was revised, during the planning stage of the evaluation, leading to a situation of co-occurrence between the project's implementation, its management (and organisational management), and its EP. This situation appeared particularly visible, in a period when the funder's expectations had been revised and the focus on EP had increased, as Glenn reported: "*we have been told by GMA: "Look, you've got to improve your evaluation"*".

The change observed in EP was that as a result of its overlapping with other WAs, a much more pressured EP was reported, and observed, as Dale illustrated: *“this pressure always forces you to [evaluate]”*. Equally, Eli’s account suggested a disadvantage of overlapping for their EP: *“So, I am doing administration, finance, management, and cleaning.” Sometimes I feel that I have not been able to be in my role in some of the areas I have expertise*”. As previous literature on EP appears to not have investigated the issue and effects of overlapping in EP when co-occurring with other WAs, my study discusses change from a DSP lens, to emphasise that accidental encounters and changeable episodes tend to occur, thus practices are constantly being shared, consciously or unconsciously (Shove *et al.*, 2012). From this assumption, a situation of overlapping, as described in section 8.3.1, highlighted that in a context of co-occurrence, each episode appeared to have its own specificity, and thus they overlapped differently. As a result, a categorisation of the overlapping of EP with other WAs over time is proposed, to argue that the effects (of co-occurrence) also change over time; possibly influencing how evaluation is practised in a global and changing community, as it is DE. Further implications for DE practitioners regarding their perception of evaluation and future practice are highlighted in section 9.4.3. Table 8.4 illustrates the categories of overlapping observed in two PCEs, and an explanation of each type follows.

Practice-change episode (PCE)	Stage in evaluation timeline	Instances of co-occurrence observed	Categories of overlapping
Funders’ different approaches to evaluation	Planning	Evaluation practice, implementation and management	1. Overlap of multiple practitioners and EP with other WAs
NDEC’s extraordinary meeting	Reporting	Evaluation practice and organisational management	2. Overlap of EP with other WAs by a single practitioner

Table 8.4: Categorisation of overlapping in a co-occurrence working environment

1) Overlap of EP with other WAs by multiple practitioners

The practice of evaluation, which has overlapped with other WAs across the timeline, has been conducted collectively by multiple practitioners, depending on a variety of external and internal circumstances; suggesting that EP is a collaborative and collective practice. As a multitude of other WAs dwell in a DE context, evaluation is practised alongside management, implementation, and negotiation/decision-making activities (section 8.3.2 and 8.3.3).

2) Overlap of EP with other WAs by a single practitioner

EP came under pressure when other WAs such as organisational management emerged to deal with the strategic future of the organisation. In this situation, most practitioners ignored the evaluation and other WAs to mainly focus on this management activity; while a single practitioner, Jesse, overlapped the EP, with other activities until the end of the project. This type of overlapping suggests that whilst, evaluation did not seem to be a priority for most, for a single practitioner - Jesse - it has a different meaning attached, possibly regarding the responsibility to meet the funder's expectations and deadline.

Although various instances of COoEP with other WAs were observed across the project timeline, this research illustrated two categories of overlapping to emphasise their trajectory across two 'moments of evaluation' (McCluskey, 2011). I refer to these as 'evaluation practice pre- YP' and 'evaluation practice post-YP', where the former was a relaxed approach to performance, less strict in demonstrating evidence; and the latter, a stricter performance-based approach to evaluation, managed by an external consultancy on behalf of DfID. Practitioner's mind-set (meaning) and skills (competence) during the evaluation practice, pre-YP, seemed to have attributed a minor role to evaluation; hence, shaping participants'

perception of evaluation. In thinking that evaluation can have a minor role, practitioners tend ‘to become what they think about’ (Nightingale, 2013), in this case practising evaluation as having a secondary role. These accounts of pre-YP evaluation practice were captured by the retrospective nature of my research that allowed participants to identify important events in the past that might have shaped their perception of the evaluation (see section 5.3.2.1; retrospective design). For instance, the example of Jesse’s involvement in an unsuccessful project that generated a successful, but apparently fabricated evaluation, illustrated how that situation may have shaped their perception of evaluation (see section 7.5.3; scepticism in evaluation).

The findings reported in section 8.3, combined with my observations, showed how the post-YP evaluation practice overlapped with other WAs such as, implementation, management, decision-making, and negotiation, but this overlap has also changed over time, as illustrated in Table 8.4 above, through two categories of overlapping. This study indicated that practitioners’ attribution of meaning to evaluation seemed deeply influenced by an initial overlap of practices through the funders’ revised set of expectations for evaluation, culminating with an overlap of EP with other WAs, by a single practitioner during an extraordinary meeting. It also underlines that there are other (new) WAs that can restrict or replace existing practices (Shove *et al.*, 2012).

8.4.2 Dominance

Dominance is another form of COoEP and WAs that can be observed when these are collaborating or competing for resources and attention. However, DSP literature highlighted that it does not always seem easy to depict whether the interactions between practices are collaborative or competitive. Furthermore, in collaborating or competing, some practices and WAs can obtain dominant status (Shove *et al.*, 2012), as illustrated when NDEC’s

organisational management co-occurred with the reporting stage of EP (section 8.3.2; an extraordinary meeting). This PCE suggested that when activities or practices co-occur, the internal and external pressure of the context is influential, supporting the DSP argument that a certain practice needs the existence of other practices (or WAs) to be enacted. During this episode, the organisational management, as well as staff working issues appeared to dominate the limited space, attention, and available resources for practitioners to keep practising evaluation – so a situation of dominance was observed during the last months of the evaluation. In competing for time, space, attention, and other potential resources, the activity of organisational management appeared to have obtained dominant status.

In addition, this research found that a commonality of the three PCEs reported in section 8.3 was the reference to anxiety and frustration generated by the process of EP. In this respect, literature on EP noted that evaluators tend to develop signs of anxiety about evaluation when tension, misunderstanding, and confusion emerged around its practice (Donaldson *et al.*, 2002). My research showed how practitioners may also develop signs of anxiety and frustration, which increase in a situation of co-occurrence. This thesis expands the literature on EP, by adding practitioners to the list of anxious members in the evaluation community.

8.5 Summary and conclusions to the chapter

This chapter reported and discussed the theme of co-occurrence, by analysing the intersection of EP, with other WAs co-occurring in a DE context. Findings and discussion were spanned across four sections as summarised in Table 8.5 (overleaf).

Issue of co-occurrence of evaluation practice with other working activities (COoEP with other WAs)	Findings emerged	Location
Evaluation practice (EP)	Evaluation is practised as a performance measurement activity in a DE context; which from participants' view implied a difficult practice that they are not able to pursue without external support. Yet, they have to conduct other working activities as required internally by their organisation.	Sections 8.2.1
Working activities (WAs)	In a DE workplace, several working activities occurred such as, reporting, regulatory, monitoring and management. These WAs co-occur with the evaluation practice, leading to urgency, anxiety, and pressure.	Section 8.2.2
PCEs of COoEP and WAs	<p>1) Change in evaluation approach: implementation and management WAs co-occurred with the practice of evaluation planning; because GMA demanded evaluation plans to be submitted from the onset; resulting in experiences of tension and pressure.</p> <p>2) NDEC's extraordinary meeting: organisational management WA co-occurred with the reporting of evaluation;</p> <p>3) NDEC's change of office: emphasis on organisational management and decision-making/negotiation WAs with the subsequent implication of a delayed completion of the final evaluation report.</p>	Section 8.3
The change observed in evaluation practice	<p>1) Overlapping: two PCEs illustrated overlapping of EP and WAs, two categories of overlapping emerged (overlap of EP with other WAs by multiple practitioners, and overlap of EP with other WAs by a single practitioner)</p> <p>2) Dominance: organisational management appeared to have obtained dominant status over EP, observed in NDEC's extraordinary meeting.</p>	Section 8.4

Table 8.5: Summary of findings on the co-occurrence of evaluation practice with other working activities

The chapter described the activities participants recognised as their evaluation practice, as well as other WAs conducted at the same time. It presented findings on COoEP with other WAs, through the depiction of three PCEs, illustrating different dynamics of co-occurrence, to argue how the project EP has changed over time. From the examination of participants' EP and other WAs, the chapter concluded that these tend to co-occur either through overlapping (section 8.4.1) or dominance (section 8.4.2). The chapter concluded that the co-occurrence by overlapping had two categories: overlapping of EP with others WAs by multiple practitioners and overlapping of EP with other WAs by a single practitioner, in the practice of evaluation. The repercussion of the COoEP with other WAs for practitioners was the realisation that DE as an unstable working environment is subject to a wide realm of changes, shaping practitioners' WAs and EP. As such, this chapter reinforced the previous argument in chapter six that the way the YP was designed influenced the way evaluation was perceived and practised over time; and the argument in this chapter is that COoEP with other WAs may influence and change how evaluation is practised in a DE working environment.

9 CONCLUSIONS

9.1 Introduction and structure of the chapter

This chapter presents the theoretical and empirical contributions of the research, alongside the implications of the findings for the domain of evaluation practice (EP), in general, and the Development Education (DE) domain, in particular. In addition, it revisits the research goals, summarises answers to the research questions, presents the implications and limitations of the research and identifies some areas for future research. Finally, it revisits my journey, introduced in chapter one, emphasising the strategy adopted to write this thesis – writing retreat as a daily practice. The structure of the chapter is represented in Figure 9.1 (overleaf).

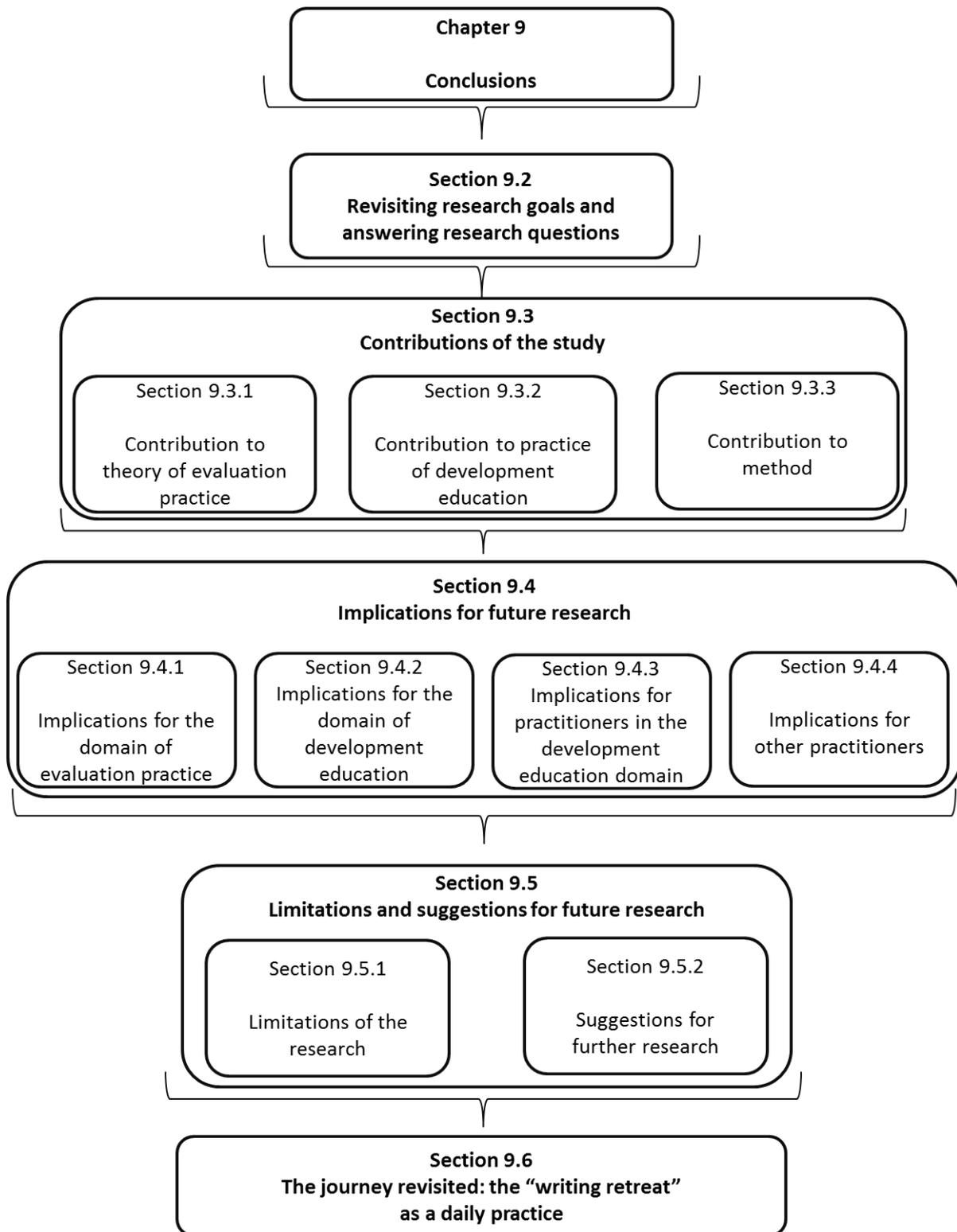


Figure 9.1: Structure of Chapter Nine

9.2 Revisiting research goals and answering research questions

The main research goal of this study was to explore and gain further understanding on how DE organisations evaluate their social interventions. Chapter two has discussed the main theoretical foundations of EP, alongside the current research debates, and in chapter three, a social practice theoretical (SPT) orientation was discussed as an appropriate framework to address the issue of how DE organisations practise their evaluation of social interventions. The focus of the study was therefore situated at the intersection of two frameworks: studies on EP informed by SPT; and studies on the dynamics of social practice (DSP), in the context of everyday life and how it changes. The first framework outlines how the reconceptualisation of evaluation as a social practice enhances current EP in some domains of social sciences (HE; e-learning; EU structural funds). This approach supports a perspective that evaluation is social, and occurs within the routinised socialisation of the everyday life (e.g. a casual conversation with colleagues can be an evaluation moment). The second framework provides an overview of how the trajectories of practice captured over time can enhance its understanding, through reflecting on how EP is carried across disciplines, practitioners or contexts. One of the key assumptions of a dynamic view of social practice is that the elements of practice are influential to the extent that a practice persists or changes over time. This DSP framework apparently has yet to be applied to studies in EP, which is one of the gaps this research fills. The key issues that emerged from the theoretical discussion of these frameworks raised further enquiry on how evaluation is practised in a DE context and how it changes over time, especially when addressing the study from a longitudinal perspective. As a result, this study has investigated the influences that shape EP, in a DE context, and how they changed over time (for this purpose, the rationale for using a DSP framework was proposed).

One of my initial concerns has emerged from the assumption, in the literature, that non-profit organisations (NPOs) in the DE sector conduct poor evaluations (Carman, 2007; 2009), which led this study to ask the following overarching question: **How do Development Education organisations evaluate their social interventions?** This question was then narrowed to three research questions (Figure 9.1), examining what do the case organisation's stakeholders do (RQ 1); the influences shaping their EP (RQ 2), and how it changed over time (RQ 3). Drawing on a constructionist ontology, a single in-depth case study design (Simons, 2009) was conducted to longitudinally approach the field work. After ethnographic data collection, a thematic analysis was undertaken using a DSP framework (Shove *et al.*, 2012), and three main themes inductively emerged and were further explored – participation, power relationships, and co-occurrence of EP with other working activities (WAs). The empirical discussion was structured around these themes and each chapter has revealed rich insights that combined offered the means to answer the research questions. Figure 9.1 gives concise answers to the research questions connecting with the previous summary of findings. The red arrows in the figure represent unexpected findings that my dataset answered, without having been initially asked.

How do Development Education organisations evaluate their social interventions?

RQ1: What are the NDECs stakeholders doing when they are evaluating the YP?

RQ2: What are the influences that shape the NDEC's evaluation practice (particularly the YP)?

RQ3: How has evaluation practice changed across the YPET?

Performance-measurement activities; formal program evaluation and design logic models.

The role of the context, the inconsistent level of participation of practitioners in the evaluation practice, the configurations of power relationships, and the co-occurrence of evaluation practice with other working activities influenced NDEC's evaluation practice. **Each influence changed over time across the YPET.** →

The YP's practice of evaluation changed according to the change of their influences. Most visible were the effects of adaption, modification, competition, dominance and overlapping.

There is **co-occurrence** of evaluation practice with other working activities

How? Concurrent practices in a DE context are:

1. Reporting ↓
2. Regulatory ↓
3. Monitoring ↓
4. Management

Evaluation is a collaborative and collective endeavour

There is co-occurrence of competences

1. Competence led by performance
2. Competence led by practice

Dominance effect of the former

Research highlighted the benefits of a **retrospective reflection** on evaluation practice ↓

The role of context (external and internal)

1. External (balanced context and imbalanced context)
2. Internal (organisational constraints + practitioners' past experiences)

The inconsistent level of participation of practitioners in evaluation

A. Typology of practitioners' engagement →

1. Disengagement (importance of project design)
2. Forcing
3. Resistance
4. Coping
5. Collaboration

B. The paradox of participation:

1. Practitioners were not encouraged to participate in the project's design; (design stage) Vs.
2. Practitioners were forced to participate in the project's evaluation (subsequent stages).

Novel configurations of power relationships

1. Leadership-led power
2. Expected power relationship
3. Competence-led power relationship

Consequences of power relationships in a funder-recipient based context

1. Legitimised dishonesty
2. Fear of judgement
3. Scepticism

The co-occurrence of evaluation practice with other working activities

The CHANGE in the role of context (external)

1. Balanced – imbalanced context (**Modification/Stabilisation**)
2. Imbalanced-balanced context (**Dominance**)

The CHANGE across the inconsistent participation of practitioners in evaluation

- A. Typology of practitioners' engagement **Modification**
- B. The paradox of participation: **Adaptation** ↔ **Modification**

The CHANGE in the configuration of power relationships over time

1. Leadership-led power **Modification**
2. Expected power relationship **Competition**
3. Competence-led power relationship

The CHANGE in the co-occurrence of evaluation practice with other working activities

1. Evaluation + Organisational management + Implementation = **Overlapping** (planning stage)
2. Evaluation+ Organisational management = **Dominance** (reporting stage)
3. Evaluation + Management = **Overlapping** (reporting stage)

Figure 9.2: Research questions answered

In answering RQ1 - What are the Northwest Development Education Centre (NDEC) stakeholders doing when they are evaluating the youth project (YP)? The study has found that the NDEC stakeholders conduct performance-measurement activities, formal program evaluation and design logic models, as their practice of evaluation. However, this study has also shown that these tasks perceived by participants as evaluation, co-occurred with other WAs, such as management, implementation, negotiation, and decision-making (not only within the project under investigation, but within the case organisation). The research has revealed that evaluation is a collaborative and collective endeavour and that co-occurrence with other WAs challenges practitioners' participation in evaluation over time. For instance, when new individuals join the evaluation during its ongoing course, its practice tends to alter, because they bring along other habits and routines – their way of doing and thinking – which interacts with the ongoing EP.

The study concludes that although the meaning attached to evaluation, by DE practitioners, is performance-based, they appear willing to respond to a call for a practice-based approach to evaluation, reflective in nature, in which an informal talk, an event, a picture or a poem can be considered 'better evidence', rather than a 'formal evaluation' (Parry-Crooke, 2014). The presentation of findings and discussion in chapter eight, allowed me to answer RQ.1 and RQ. 3, and to avoid repetition. The intrinsic interconnectedness of these reported findings accounts, I argue, for the robustness of my research method, particularly regarding the longitudinal immersion in the field that gave me various possibilities to connect and make sense of these findings.

In answering RQ2 - What are the influences that shape NDEC's evaluation practice, particularly the YP? This research has found that the influence of the context, an inconsistent level of participation of practitioners in EP, and novel configurations of power relationships

strongly shaped the case organisation and the project EP. The role of the context appeared as a cross-cutting influence of how the issues of participation, power, and co-occurrence of practices shaped EP. Explicitly this influence has emerged as a funding-recipient dimension, with an external political strand of setting the evaluation goals (funders' performance-based approach), alongside a situated, internal strand that highlights the constraints faced by the DE domain, when evaluating public funded-projects. These constraints were also perceived to influence practitioners' past experiences on their current EP. The research has shown that, unless the nature of DE organisations' financial sustainability intentionally changes, the rules of the funding context are likely to remain power-led; regardless of the dynamics of change observed (thus, the role of context is likely to remain influential).

The research has also found that an inconsistent level of participation in practitioners' EP influenced their practice of evaluation at a micro (YP) and meso (NDCE) level; responding to this aspect of RQ.2. However, this finding provides fresh insights through a typology of practitioners' engagement in EP, adding the 'how' to RQ.2: How has an inconsistent level of practitioners' participation (identified as influence) influenced the organisation and project's EP? The research has highlighted how practitioners are deliberately left out of the evaluation (project design stage) and, at other times, forced to engage in it (planning stage), particularly, when funders' deadlines are approaching. This study concludes that an inconsistent level of participation in EP appears over time; which emerges from a typology of practitioners' engagement in the evaluation (discussed in section 6.4) and from a paradox of change observed in their participation in EP over time (discussed in section 6.5). In addition, the thesis sheds light on the barriers to practitioners' participation in EP, such as leadership, funding, and the co-occurrence of EP with other WAs, and how these tend to subtly become entangled in the speediness of the daily routine.

The study has revealed that the presence of power particularly manifested through novel configurations of power relationships – leadership-led power; expected power and competence-led power - shapes the case organisation and the project EP. Findings have identified that participants evaluate the project within a circular shifting relationship, in which a balanced context shifts to an imbalanced, and re-occurrence occurs (section 7.2; the context of funder-recipient based evaluation). These findings intersect with findings of power relationships, as this study suggested that within a balanced context, configurations of an expected power relationship appeared (funders and organisations have clear expectations about the process and guidelines for evaluation, so their power is balanced); whereas within an imbalanced context, configurations of leadership and competence-led power relationship are more likely (because the role of the leader and the element of competence appeared prominent at some stages of evaluation, resulting in an imbalance of power; see section 7.3). From these findings around power relationships, a sense of legitimised dishonesty, judgement and scepticism appears as a possible consequence of the influence of power when practising evaluation (discussed in section 7.5). This study has found that practitioners, when confronted with an established funding-dependent context, tend to tweak and adapt their practice in a slightly dishonest behaviour, which may become legitimised, within a particular workplace or context. This conclusion is likely to be transferrable to other NPO domains with a similar funder-recipient based context (see section 9.4.2; implications for the domain of DE).

In answering RQ.3 - How has evaluation practice changed across the YP evaluation timeline?, this research has revealed that different effects of change arose across the YP evaluation timeline (e.g. adaptation, modification, competition, overlapping, and dominance). The study has shown that when the NDECs' practitioners appeared disengaged, during the project design, their expectation of, possibly, no longer participating was adapted, due to a

‘forced invitation’ to engage in and participate in the subsequent stages of the evaluation – a modified practice emerged. As a result, I argued that their perception and practice of evaluation has changed. The effect of modification was illustrated by the change that occurred across a typology of engagement where practitioners’ participation was modified over time (disengagement, forcing, resistance, coping, and collaboration). Hence, the research concludes that subsequent effects of modification have generated a paradox of participation (disengagement/forcing) that in, turn, revealed how EP may change or persist over time.

Another effect of change that emerged from the research was competition. A competing two-way dynamic between the elements of funding and competence emerged (illustrated in Figure 9.1, by a double red arrow in the RQ.3 column). In one way, the dominance of funding appears to overshadow the focus on practitioners’ competence (see section 6.2.1; funding dependency versus competence), whereas in the other way, the DEC partner’s competence seems to compete with the dominance given to the funding (see section 7.3.3; competence-led PR). This conclusion illuminates how different effects of change – modification and competition – can intersect across configurations of power relationships (leadership, expected, and competence-led power).

The research has also found that EP appeared to have changed over time, due to its overlapping with other WAs, as well as the dominant effect of organisational management over EP. The relevance of evaluation as a social practice is supported by the current findings, in the way that a practice-based evaluation maintains that the process of attributing worth and value is social and relational (Saunders *et al.*, 2011), in which change is bound to occur.

The three main themes of this research and their respective findings provide insights for the influences of EP, and how it is enacted, as well as for the evaluation practised in a DE

domain. The combination of my findings allowed me to suggest that, in a DE context, evaluation is not the central practice of practitioners' work – other WAs occupied their working environment – which corroborates past literature recognising the existence of other practices alongside evaluation (Saunders, 2000). Nonetheless, this study went beyond this recognition by having explored how the intersection of other WAs may have changed EP, and with what effects. As a way of summarising this section, Figure 9.2 (overleaf) offers an overview of how EP changed over time. The next section elaborates on the theoretical, empirical, and methodological contributions of this study.

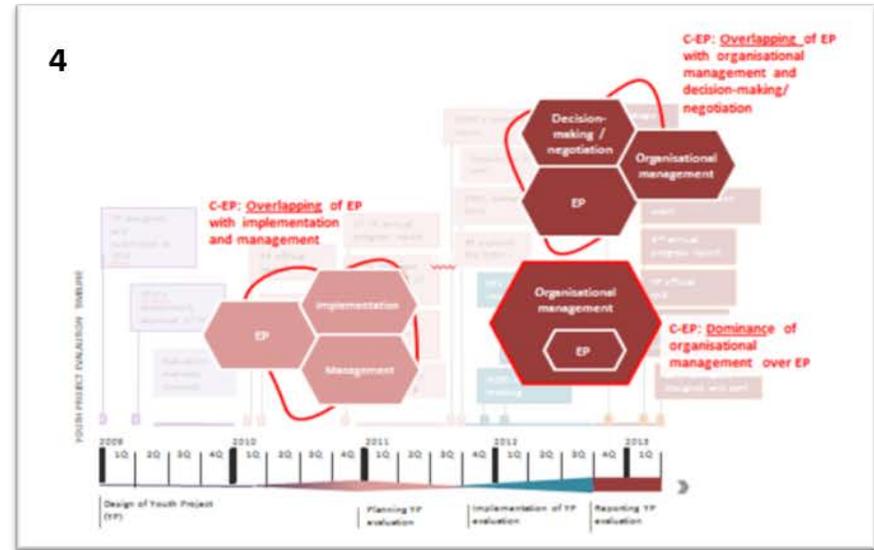
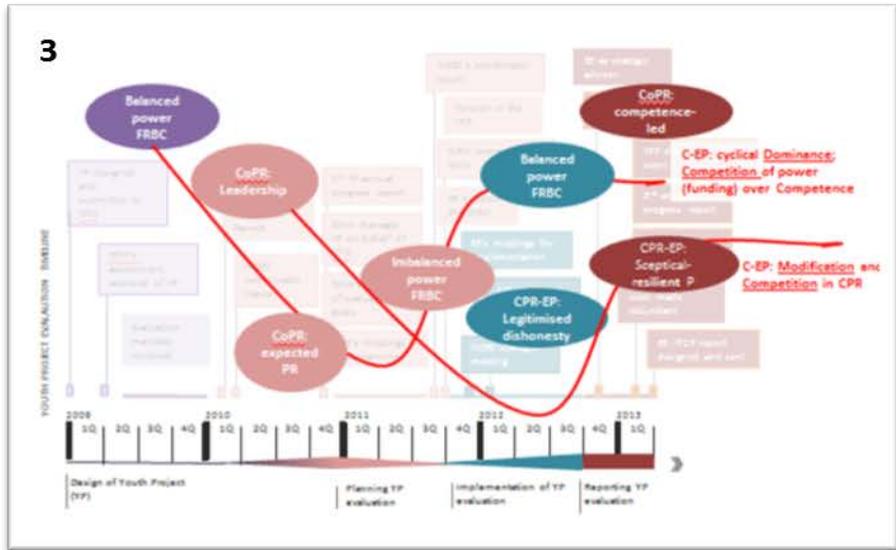
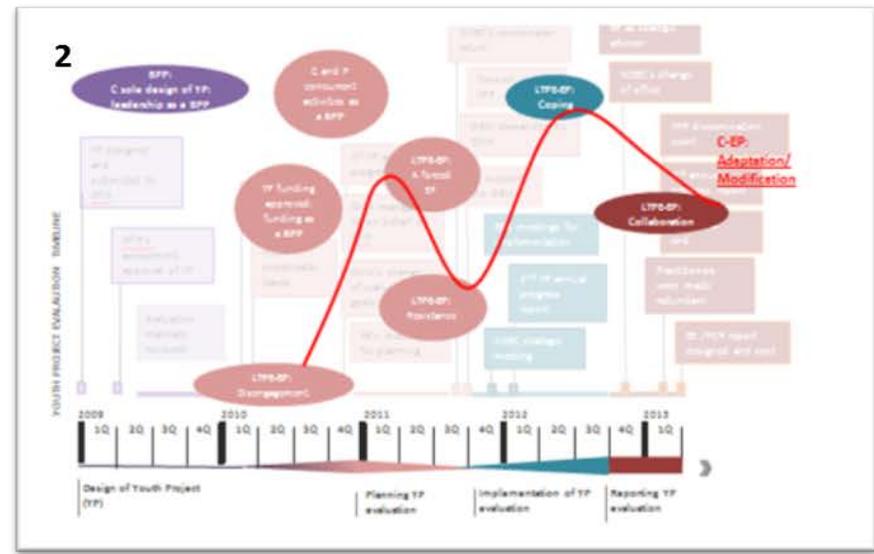
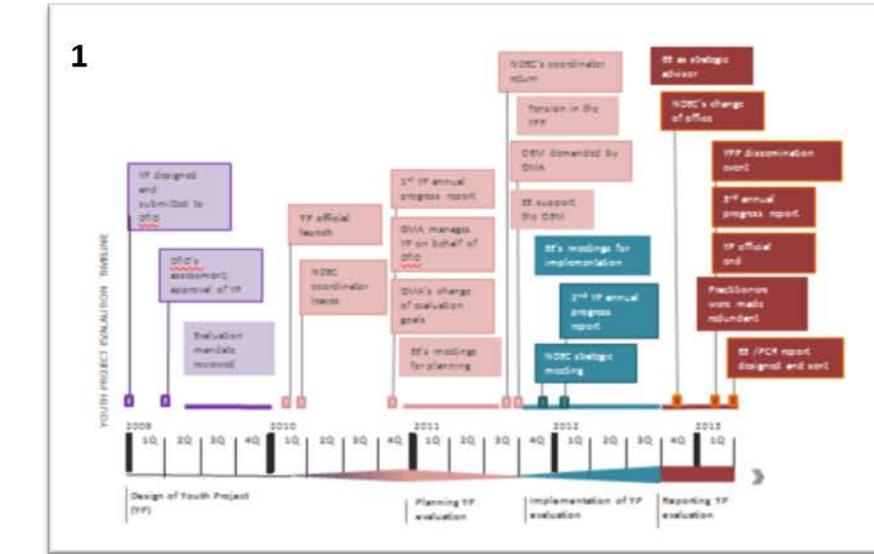


Figure 9.3 Overview of how evaluation practice changed across the YPET

9.3 Contributions of the study

9.3.1 Contribution to theory of evaluation practice

The original contribution of this thesis is to extend the theory of EP informed by SPT (Saunders, 2000; Reckwitz, 2002; Saunders *et al.*, 2011), through the use of an advanced DSP framework (Shove, 2003, 2007; Shove *et al.*, 2012), to explore the everyday life of evaluation and how it changes over time. The key strengths of this study are its long duration of observation in the field to develop current understanding of how evaluation is practised, filling the theoretical gap in the literature on how EP changes over time and with what effects (Henry and Mark, 2003; Saunders *et al.*, 2005). By drawing on Shove's DSP (2003, 2007; 2012), this research makes the following original contributions to the theory of EP (Saunders, 2000; Saunders *et al.*, 2011; Saunders, 2012).

- **Extends the theory by applying the dynamics of social practice to research evaluation practice**

This study adds to the literature through the application of a DSP lens (Shove, 2003, 2007; Shove *et al.*, 2012) in researching EP. Specifically, it contributes to the current understanding of how evaluation is practised in small NPOs of the social domain by: 1) adding depth to the current knowledge about participation of practitioners in evaluation processes; 2) proposing novel configurations of power when practising evaluation; 3) extending present knowledge on the effects of the co-occurrence of EP with other WAs.

1. Inconsistent level of participation of practitioners in evaluation practice

This study adds precision to the debate of participation in EP (Abma, 2000; Greene, 1988; Simons and McCormack, 2007; McCluskey, 2011; Simons *et al.*, 2003; Stake, 2004), by highlighting an inconsistent level of participation of practitioners in evaluation. Whilst

previous studies on EP implicitly assumed a constant (and deliberate) participation of practitioners (Abma, 2000; McCluskey, 2011), my study has revealed its inconsistency, illustrated through a typology of practitioners' engagement in evaluation. Consequently, and building on previous literature on small-size NPOs (Greene, 1998; Bourn, 2014), this thesis advocates that the empirical findings observed in this study provide a new understanding of the issue of participation in EP, and therefore may occur in other similar NPO domains.

2. Importance of an internal configuration of power shaping practitioners' evaluation practice: leadership-led power

This research contributes to current knowledge by adding three novel configurations of power relationships (leadership-led power; expected and competence-led power). Explicitly, whilst recent research mainly emphasised the established power relationship from funders to NPOs (Atjonen, 2015); this thesis adds a contribution of an internal layer of power, revealed by coordinators (as leaders) to their team of practitioners.

3. Awareness raised for the co-occurrence of evaluation practice with other working activities, and its effects of change

This work contributes to existing knowledge on EP (Saunders *et al.*, 2011), by providing fresh insights on how the theme of co-occurrence of EP with other WAs generate change in the former. The empirical findings in this study highlighted how some effects of change (adaptation; modification, competition, overlapping, and dominance) can shape EP over time. Specifically, the study proposes that EP predominantly overlaps with other WAs, increasing pressure in a DE working environment, as well as on the practitioners' fragile sense of their competence in evaluation. These effects are likely to change how individuals perceive and practise evaluation. While previous literature on EP has recognised the importance of co-occurrence for the practice of evaluation (Saunders *et al.*, 2011), to my knowledge, the

connection between that finding and consequent effects of change in EP have not yet been substantially explored. Therefore, this research extends previous knowledge, by suggesting that evaluation is a collaborative and collective practice, with multiple other endeavours co-occurring.

- **Extends the theory by applying the dynamics of social practice into a novel research domain of Development Education evaluation**

This thesis made an original addition to the literature on EP, by using DSP, as an advanced social practice framework, to research EP in a novel domain of social sciences – Development Education. In doing so, it contributes to an in-depth understanding of how EP changes over time, by: 1) revealing the change occurring on the influence of the role of DE context, and 2) adding depth to the current knowledge about practitioners’ engagement in evaluation and its change over time, and 3) expanding the current knowledge about the role of practitioner in a NPO domain, as an essential carrier of EP.

- 1) **Emphasis on how evaluation practice changes in accordance with the change of their influences, specifically, the role of the DE context**

My study found that when the context changes, from a balanced to and imbalanced one, regarding the expected conditions for evaluation, its practice is modified, whereas change occurs by dominance within an imbalanced context, in which the conditions for evaluation are unexpected (Figure 7.2; balanced and imbalanced power across YPET). Previous studies on EP in NPOs highlighted the importance of researching the role of context (Christie, 2003; Rossi *et al.*, 2005; Saunders *et al.*, 2005; Simons, 2009; Vo and Christie, 2015), particularly regarding the expected dependency from funders’ goals; however, the issue of how change in context influences change in EP appeared underestimated. The funding-dependent context

that this study has identified, therefore, assists in the understanding of how contextual change shapes change in EP.

2) Importance of change across the typology of practitioners' engagement in evaluation practice

This study extends past research on EP by adding the dynamics of change to the exploration of practitioners' engagement in evaluation. Previous studies assumed that the more practitioners engage in and participate in evaluation; the more they are likely to learn in and from it (Wenger, 1998; Saunders *et al.*, 2011). From a DSP lens, the thesis found various types of practitioners' engagement in evaluations that were modifying over time: from a state of disengagement, to forced participation, resistance, coping and then collaboration. Each modification changed practitioners' engagement with evaluation, which meant that their participation was irregular, shaped their EP, and in turn influenced others' perception (and practice) of evaluation.

3) The under-researched role of practitioners in NPOs, as an essential carrier of evaluation practice

Another key strength of this study is the emphasis on the role of practitioner in NPO domains, as an essential carrier of EP (Shove *et al.*, 2012), by emphasising their intention to learn about and from evaluation, to create alternative tools to demonstrate evidence, and to enquire about unresolved issues within their EP. This intention was observed in my study, and in some situations explicitly shared, particularly when contextual and power structures pointed to other directions, other than the learning, the creativity, and the critical thinking in understanding and responding to funders' request for the evaluation (in a meaningful way for practitioners too). The role of practitioner, as an essential carrier of EP, and driver of their own destiny finds common ground with SPT and democratic approaches to evaluation; but

goes beyond these, touching a social justice stance of evaluation (Mertens, 2005). Notable exceptions on the under-researched role of practitioner in EP (Greene, 1988; Taut and Brauns, 2003; Seppanen-Jarvela, 2004) have called for more contributions on that matter; thus, my study represents a relevant addition to the literature on EP, by advocating the role of practitioner, as an essential carrier of a meaningful practice of evaluation.

9.3.2 Contribution to practice of development education

This thesis offers an empirical contribution to the under-researched domain of DE (Bourn, 2011, 2014), by extending current knowledge on how evaluation is practised. Current literature in the DE domain called to further investigation on evaluation, particularly, on the need to “*place practice at the heart of the evaluation*” (Bourn, 2014, p.50).

This study qualitatively explored EP within a small NPO, in the a-political domain of DE, to fill the analytical gap as past empirical research has mainly analysed large and medium-size NPOs, within popular and politicised domains of social sciences, using quantitative and mixed-methods (Carman, 2007, 2009; Carman and Fredericks, 2010). Consequently, some questions remained unanswered, particularly, regarding issues influencing change in practitioners’ EP. Moreover, there is scant literature on DE evaluation, with the notable exception of Bourn (2011; 2014), who argued for the need of further research reinforcing the importance of a practice-based approach to evaluation. By drawing on Shove’s work (2003, 2007; 2012) to explore how evaluation was practised in a DE domain, this research makes the following contributions to that domain.

- **Adds depth to the under-researched domain of Development Education**

The study extends current knowledge on how evaluation is practised and changed over time by: 1) adding a SPT view of evaluation that humanises practitioners’ EP, in responding to

fundings' requests; 2) expanding the present knowledge of a DE working environment; and 3) extending previous literature on power relationships in NPOs.

1) Proposes a SPT view of evaluation that humanises practitioners' EP

My study advances the current knowledge of evaluation in DE (McCollum and Bourn, 2001; Bourn, 2014, 2011), by revealing that a humanised view of evaluation may empower practitioners in their EP, if there is willingness to gradually step out of their comfort zones (as the research suggested there is). This is an exciting critical momentum¹⁷ to advance and establish a social practice alternative of evaluation that can humanise how DE practitioners, respond to the current evaluation requested by funders. Whilst previous attempts were made to bring a humanised view to evaluation (research and practice) through the shift of focus from evaluation theory towards participation of practitioners within the evaluation processes (Greene, 1999; Mertens, 2001; Guba *et al.*, 2011; Simons, 2015), the level of their participation in evaluation seemed inconsistent, especially in a DE domain.

Despite this inconsistent level of practitioners' participation in EP and the configurations of power relationships, my research has emphasised that the current evaluation approach requested by funders is far from democratic terms of what practitioners do, think, talk, decide, as well as in how they engage (or not) in the politics of EP (e.g. negotiation with funders on alternative ways of reporting and authenticity about what does not work).

¹⁷ After 2015, International Year of Evaluation.

2) Expands DE current knowledge by placing individuals' practice at the heart of evaluation

This research found that a DE working environment, and possibly other NPOs generally, has various particularities that shape practitioners' EP, such as an unpredictable policy due to changes in government policies; funding-dependency; an unstable structure of human resources; and a multitude of WAs co-occurring with EP. The study has gone towards enhancing our understanding of a daily routine of EP, in an otherwise a-political DE setting, caught in the politics of evaluation (Simons, 1987), which practitioners thought as being the 'others', but realised were also 'theirs'. In this vein, my study concludes that these particularities are constantly changing over time, so that practitioners may accommodate some (as they did through adaption, modification, competition, overlapping, and dominance) or resist others (as they also appeared to have done).

3) The importance of protecting practitioners' authenticity, challenging established structures of power

The study has found that DE practitioners seek authenticity in their daily routines of evaluation, and when that is threatened by the various structures of power, the sense of lacking confidence in their competence, and not being good enough, arises. Also, it suggests that practitioners have intentionally decided to work in a DE context, possibly because of a strong belief in a social cause for the greater good. As a result, they appeared reluctant to the massive bureaucracy - power-led relationships and a funding-dependent context - inherent in making that good. The thesis extends previous literature on power relationships in NPOs (Abma, 2000; Turner, 2005; Alaimo, 2008; Atjonen, 2015), by adding fresh insights into the intersection of practitioners' competence with power within a DE organisational context.

9.3.3 Contribution to method

- **Use of the timeline to display findings**

The research makes a methodological contribution that emerges from the choice of the timeline as the main tool to display findings (Shove *et al.*, 2012). As in other methods, for example, life course analysis (Giele and Elder, 1998), the timeline was used to collect and organise data from participants' life stories, but neither to represent how findings (interpreted data) influenced the research topic (evaluation practice) nor to display the trajectory of that topic and its change over time. However, the design of the timeline in itself evolved from various trials, and it has my presence, as a researcher, within it. From the form and shape of the evaluation stages, to the colour coding of findings, my influence was explicit. An extension of this methodological contribution is the application of this constructed timeline to a novel domain of DE; whether the findings have possible implications for the DSP framework; and to what extent this framework is challenged by its application to a DE domain (section 9.6.2; suggestions for further research).

9.4 Implications for future research

9.4.1 Implications for the domain of evaluation practice

The research reveals the potential of a SPT approach to evaluate social interventions of small NPOs through alternative tools and reflection, which raises implications for the conceptualisation of practitioners' competence, and the role of reflection in their EP. Regarding competence, recent research on EP acknowledged, accepted, and promoted creative talented practitioners (and evaluators) (Simons, 2015) who can act as socially just, challenging established structures of power, not in an oppressive and harmful way, but from the assumption that if practitioners do not let funders know what works and what does not,

their competence and funders' expectations will not match, especially as these tend to change over time.

- **Transferability of findings**

The study indicated the necessity to develop more knowledge about how participation, power relationships and co-occurrence of EP with other WAs influence EP in other domains of social sciences. Throughout the journey, I had a sense of this implication in conversations with other peers who have described connections with their domains. Additionally, during a seminar given in the Department of Educational Research, Lancaster University, some colleagues identified intersections between the main findings of my study and their experiences of the evaluation, some in Higher Education, and others in various NPO domains (section 5.3.8; ensuring data quality). Moreover, from the insights received, the particularity of co-occurrence of EP with other working activities, as a threat to practitioners' participation in evaluation, appeared transferable to other similar settings.

- **Concept generalisation**

Following Simons' (2009) procedures, one possible way of generalising from a case study is through concept generalisation, which means that a concept arising from the findings can be transferable and applicable to other contexts, regardless of the specific instances or similarities. For example, this happened with the finding, and concept, of the co-occurrence of EP with other working activities; which this case study revealed can influence the practice of evaluation over time. This concept may well be applicable to and observed in other contexts and practices. The research highlighted particular instances generated by this finding, allowing the reader to make similar or contrasting connections with their domain, and this is preferred in the realm of a naturalistic generalisation (Stake, 1995).

- **In-depth particularisation**

Simons (2015) argued that it is possible to generalise from a single case, if the paradox of understanding the universal by examining the particularities of the case was accepted. Her argument goes beyond this paradox, and advocates the importance of a single case for policy making, because it allows policy makers to understand complex situations, only possible to be seen if explored in-depth. The in-depth particularisation of my case study allows, for example, the reader to realise the significance of practitioners' lack of confidence in their competence, or an instance where a dishonest behaviour becomes legitimate. Possibly these issues (that the literature has only superficially mentioned in a DE context) would not have been seen, had it not been explored in-depth with multiple observational interactions.

9.4.2 Implications for the domain of development education

- **An a-political but power-led evaluation practice**

One implication for this research is whether DE NPOs' actions are consistent with their a-political nature, as it is communicated and perceived. One of the criteria for sampling was the assumption about NPOs' a-political affiliations, and so NDEC appeared as an information rich-case. However, as the research evolved, the finding about power relationships, combined with an in-depth account of a funding-dependent context, flagged up the possible implication of ambiguity in the way the case organisation faced this a-political dimension. This means that awareness, as well as reflection is needed on whether a NPO's passive attitude to funders' political decisions can be perceived as consonant with their (politicised) practice of evaluation, thus contradicting NPOs' a-political aim. For instance, by not taking action and not actively saying what could be improved and how difficult some requests were, NPOs may be (unconsciously) collaborating and perpetuating a politicised practice of evaluation.

Therefore, this study has implications for the practice and politics of evaluation within the DE domain.

9.4.3 Implications for practitioners in the development education domain

This thesis has various implications for practitioners' future practice of evaluation, from which I outline four. The first concerns the need of placing their practice at the centre, and for that my study emphasises the need to encourage reflection (e.g. time, predisposition, and enthusiasm) and to discover what makes their EP distinctive. A second insight suggests competence needs to be reframed in order to understand the challenges faced by practitioners. A direct implication for DE practitioners concerns the attention given to their competence to evaluate, through recruitment, training, and learning, within a context where only funding seems to matter. A third implication alerts to the excessive dependency on external funding, by pointing to the reconfigured nature of power relationships developed in a DE domain. The perception of funding, as a barrier to participation, may increase practitioners' resistance to engage in future evaluations. Finally, the challenge in understanding the value of evaluation (Parry-Crooke, 2014) when its fundamentals (rationale, theory, purpose) have not been taught at any point of practitioners' journey may have future implications for their EP.

9.4.4 Implications for other practitioners

For the purpose of this section, I refer to practitioners as all individuals undertaking evaluation in a DE or other NPO domains. In order to facilitate a practice-based evaluation, I propose the translation of the evaluation fundamentals (e.g. theory, rationale and purpose) into action learning sets (ALS) (Smith, 2011) where a small group of practitioners meet up,

either as a team, with their funder¹⁸ or a facilitator over a pre-agreed period of time to address existent issues on their EP, through open questions and dialogue. Within these ALS, “questioning is the main way to help participants proceed with their problems, and learning is from reflection on actions taken” (ibid. p.20). As a result, practitioners can be at the forefront of their practice and possibly capture the change occurring on their mind-set over time, from a ‘fixed mind-set’ to a ‘growth mind-set’ (Dweck, 2006). The former relies on the idea that we are either born with a mind-set or are not, and we are ‘destined’ to stay (fixed) in that mode. On the contrary, a ‘growth mind-set’ emphasises the effort needed to become a little bit better every day, by practicing it (hard) consistently, diligently and, hopefully, playfully. These ALS are more likely to instigate a ‘growth mind-set’ for NPOs practitioners, through which their future practice of evaluation may be meaningful and enjoyable.

9.5 Limitations and suggestions for future research

9.5.1 Limitations of the research

- **Cautious interpretation**

Regarding the analysis and interpretation of data to explore how participants practised evaluation over time, this study had to make a cautious interpretation on various occasions, to protect participants’ confidentiality and anonymity. For instance, coordinators’ articulations of leadership, discussed in sections 6.3.1 were an example where caution was critical, because the findings discussed were likely to be sensitive and may not represent the wider

¹⁸ This recommendation aligns with the argument presented in the seminar “What if evaluation takes place seated around the table with a glass of wine?”

<https://www.youtube.com/watch?v=kx9h80FsOSY&feature=youtu.be>

sector. To ensure that caution was taken, I relied on a descriptive account of coordinators' routines, attitudes and absences, to allow the reader to make their own inferences. Nonetheless, that was not without its complications, and in some instances it was necessary to summarise part of the evidence to protect the individual, as illustrated in the description of leadership-led power (section 7.3.1) and withhold other potentially sensitive evidence in a confidential appendix (section 7.3.3).

- **What could have been decided differently?**

I could have invited the funders to fill in the evaluation form during the first phase of interviews (Appendix 5). I had assumed that because they were already practising evaluation, as their legal duty, their views would not represent their own experience of practice, but rather a policy-related account. Although I can see where this decision came from, as the analysis evolved, I thought that it would have been interesting to 'hear'¹⁹ funders' perception of evaluation.

9.5.2 Suggestions for further research

1. The role of the language of evaluation in its practice

The findings indicate that the perception of evaluation is a critical feature of its practice; however, a key dimension of EP, widely recognised in the literature, is the language of evaluation (Scriven, 1991; Patton, 2000; Madison, 2010). This thesis triggers the following questions as futures avenues for research: 1) How does a specific terminology shape the perception of evaluation? 2) To what extent does the role of language alters the practice and meaning of evaluation?

¹⁹ As mentioned in chapter five, funders have not been observed and only interviewed by skype due to their distant locations.

2. Affectivity and trust in EP

This thesis adopted a social practice approach to research EP and, several ‘hidden’ issues came to light as relevant to a deeper understanding of the phenomena that exist in small NPOs such as, affectivity and trust in evaluation. Whilst these issues may seem sensitive to be critically explored, affectivity and trust have been so - though superficially - (Love, 1996; Schwarz and Struhkamp, 2007; Dane et al., 2012), which leaves room for further research.

3. Extended application of the framework of dynamics of social practice (DSP)

In adopting a DSP framework, an inductive decision was made to explore the main aspects that emerged from data; as a result, the framework was partially applied in the analysis of my research (see chapter three). A remaining part of the framework constitutes a potential area for further research, particularly focusing on: how access to relevant elements of EP is distributed; how instances of integration themselves transform and generate new materials, competences, and meaning; and the circuits and intersecting circuits of reproduction practices; and practice-oriented policy making (Shove *et al.*, 2009, 2012).

Also, from a DSP view, the materiality of evaluation, as an element of practice was less addressed as the other elements of meaning and competence; possibly, because the type of data sources collected made a rich in-depth account of how evaluation was practised and changed over time; and the word-number limit for this thesis was reached. Also, in observing the effects of change in EP, I mostly identified their type (e.g. adaptation, modification, competition, overlapping, and dominance), but further implications of each effect in EP constitute a possible avenue for future research.

4. The role of funders' in attributing worth to recipients' evaluation in DE or other NPOs domains requires further investigation

Three possible questions left unanswered were: 1) How is NPOs' evaluative evidence judged within funders' working environment? 2) How is NPOs' evaluative evidence communicated internally? and 3) To what extent does the appearance of more actors in a funding-recipient context shape the power relationships at stake (where consultants have to please funders; funders have to please the Government; and the Government has to please their citizens - on the assumption that they do!).

9.6 The journey revisited: the “writing retreat” as a daily practice

As the course of the PhD evolved and come to an end, I can now firmly say that my writing *practice* sustained the journey and made it enjoyable, yet challenging. A writer's journey is always unique and challenging; but to write academically, for the first time in my third language²⁰, adds some more complexity to the whole process. I knew from an early stage that my academic writing might be a potential weakness; so the strategy used to enhance my writing was discovered during a writing retreat organised by Professor Rowena Murray²¹, with whom I had the privilege to write. During a weekend, I discovered a whole new topic of writing and how I could ‘even’ enjoy the experience of writing academically in a foreign language (which meant that I had to deal with issues of confidence and anxiety). The term “*writing retreat*” means a “*space of dedicated writing time where collegiality is built*” (Murray and Newton, 2009, p.541). As a structured intervention, writing retreats have three distinctive features: “*all writing together in one room for the whole of the retreat; structuring*

²⁰ My mother tongue is Portuguese and my second language was Spanish, at the start of the PhD. Now the English language has been competing and won the second place in the ranking!

²¹ www.rowenamurray.org accessed on 13th December, 2015

retreat time as a series of fixed writing and discussion slots; and discussing writing-in-progress throughout the programme” (ibid. p.542).

Three aspects profoundly inspired me during the retreat: the first was the awareness that many different people struggle with writing (even in their native language), regardless of their position in the career ladder. Second, that it was possible to finish the PhD if the perceived ‘huge writing piece’ was split into manageable little pieces, or in Murray’s terms, *writing slots*. For that, I ‘only’ had to show up day in, day out, persistently. The image of a manageable PhD, rather than a ‘scary elephant’ arose early on in my journey, keeping my motivation going, as well as helping me to overcome some obstacles along the way. Third, the idea of planning the writing was visible through an initial sharing about one’s goals and a final debrief about one’s achievements or difficulties, followed by a peer support conversation. That was revealing to me, as I have been impressed with my own progress, since I started to deeply think and thoroughly plan my writing in that slot-oriented way. After this initial writing retreat, I was committed to replicate its structure into my daily practice and, as a result, I set up a writing group in the Department of Educational Research, at Lancaster University. The excuse for organising this group could have been interpreted selfishly, as I was committed to writing in a supportive environment, booking a different room with no internet connection; if any peers showed up that would help, otherwise, I would write anyway. Since the beginning, some colleagues have joined in and the group has been critical for some PhDs’ successful completion, and I hope for mine, too! Having this structured writing space allowed me to experience the notion of a “*routinised practice*” (Reckwitz, 2002, p.259), while I was researching about evaluation as a “*social routinised practice*” (Saunders *et al.*, 2011, p.62); in this respect it was a very fortunate coincidence. Also, having such a privileged space to write allowed me to construct an operational document, used as an ongoing learning log, which gradually changed into a ‘treasure log’

with rich data quotes, thoughts, and reflections. This ‘treasure log’ was then refined and transformed into a first draft; a first final draft; and this final thesis.

As I finalise writing this thesis, I realise that I do not really want to finalise it, because I enjoyed the *practice* of writing, so much. In revisiting the journey, I can say that the reflection on my own writing practice and how it changed over time was an unexpected, additional, and powerful outcome of this research ... I already have ideas to keep writing after the PhD.

This thesis is a construction that has been co-created, primarily by me and my research participants. As a researcher, my identity has changed over time (most by *modification* with some occasions of *competition* between identities), and was translated from mother/wife to PhD student, from practitioner/volunteer to researcher, through to becoming (I hope) a Doctor of Philosophy. This process involved the reader, the extended PhD community, my supervisor, and my family, who acted as co-creators of and in this journey. In honour of recognising this as a closing statement, I would like to express my gratitude to them and to the examiners for being part of this process of co-creation. I hope you have enjoyed reading it, as much as I enjoyed writing it.

10 LIST OF REFERENCES

- Abma, T.A. (2000) Stakeholder conflict: a case study. *Evaluation and Program Planning*. **23**(2), 199–210.
- Abma, T.A., Widdershoven, G.A.M. (2008) Evaluation and/as Social Relation. *Evaluation*. **14**(2), 209–225.
- Adams, P. (2010) *A history of global youth work*. Development Education Association.
- Akiba, D., Alkins, K. (2010) Learning: The Relationship Between a Seemingly Mundane Concept and Classroom Practices. *The Clearing House*. **83**(2), 62–67.
- Alaimo, S.P. (2008) Nonprofits and evaluation: Managing expectations from the leader's perspective. *New Directions for Evaluation*. **2008**(119), 73–92.
- Alkin, M.C. (2003) Evaluation theory and practice: Insights and new directions. *New directions for evaluation*. **2003**(97), 81–90.
- Alkin, M.C., Daillak, R., White, P. (1979) *Using evaluations: does evaluation make a difference?* 2. print. Beverly Hills: Sage Publ.
- Alkin, M.C., Ellett, F.S. (1985) Evaluation models and their development. In T. Husen & T. N. Postlethwaite, eds. *International Encyclopedia of education: Research and Studies*. Oxford, England: Pergamon Press, pp. 1760–1766.
- Alkin, M.C., House, E.R. (1992) Evaluation of Programs. In M. C. Alkin, ed. *Encyclopedia of Educational Research*. Old Tappan, N.J.: Macmillan.
- Atjonen, P. (2015) 'Your career will be over'—Power and contradictions in the work of educational evaluators. *Studies in Educational Evaluation*. **45**, 37–45.
- Ayers, T.D. (1987) Stakeholders as Partners in Evaluation: A Stakeholder-Collaborative Approach. *Evaluation and program planning*. **10**(3), 263–71.
- Bamber, V., Trowler, P., Saunders, M. (2009) *Enhancing learning, teaching, assessment and curriculum in higher education theory, cases, practices*. Maidenhead, England: McGraw-Hill : Society for Research into Higher Education & Open University Press.
- Bamberger, M., Segone, M. (2011) *How to design and manage Equity-focused evaluations*. Unicef.
- Barton, D., Tusting, K. eds. (2005) *Beyond communities of practice: language, power, and social context*. Cambridge ; N.Y: Cambridge University Press.
- Becher, T., Trowler, P. (2001) *Academic Tribes And Territories: Intellectual Enquiry and the Culture of Disciplines*. Society for Research into Higher Education & Open University Press.
- Becker, H.S. (1977) *Boys in white: student culture in medical school*. New Brunswick, N.J: Transaction Books.

Bemelmans-Videc, M.-L., Rist, R.C., Vedung, E. eds. (1998) *Carrots, sticks & sermons: policy instruments and their evaluation*. New Brunswick, N.J., U.S.A: Transaction Publishers.

Berner, M., Bronson, M. (2005) A Case Study of Program Evaluation in Local Government: Building Consensus through Collaboration. *Public Performance & Management Review*. **28**(3), 309–325.

Berry, C., Barnett, E., Hinton, R. (2015) What does learning for all mean for DFID's global education work? *International Journal of Educational Development*. **40**, 323–329.

Bhola, H.S. (2003) Social and Cultural Contexts of Educational Evaluation: A Global Perspective. In T. Kellaghan & D. L. Stufflebeam, eds. *International Handbook of Educational Evaluation*. Dordrecht: Springer Netherlands, pp. 397–415.

Blaxter, L., Hughes, C., Tight, M. (2010) *How to Research*. 4th ed. Open University Press.

Bonamy, J., Charlier, B., Saunders, M. (2004) The Evaluative Research of Complex Projects in e-Learning: The Case of the EQUCEL (e-Quality in e-Learning) Project. In *Networked Learning Conference, Lancaster, UK*.

Bourdieu, P. (1977) *Outline of a theory of practice*. Cambridge; New York: Cambridge University Press.

Bourn, D. (2008) *Development education: debates and dialogue*. London: Institute of Education, University of London.

Bourn, D. (2011) Discourses and Practices around Development Education: From Learning about Development to Critical Global Pedagogy. *Policy & Practice - A Development Education Review*. (13).

Bourn, D. (2014) *The Theory and Practice of Global Learning*. Development Education Research Centre.

Bracken, M., Bryan, A. (2010) The reflective practitioner model as a means of evaluating development education practice: Post-primary teachers' self-reflections on 'doing' development education. *Policy & Practice - A Development Education Review*. **11**, 22–41.

Brody, E., Cordes, J.J. (1999) In *Nonprofits & government: collaboration & conflict* - Elizabeth T. Boris, C. Eugene Steuerle - Google Books. Washington DC: The Urban Institute.

Brown, R.D., Newman, D.L., Rivers, L.S. (1985) An Exploratory Study of Contextual Factors as Influences on School Board Evaluation Information Needs for Decisionmaking. *Educational Evaluation and Policy Analysis*. **7**(4), 437.

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

Carman, J.G. (2007) Evaluation Practice Among Community-Based Organizations: Research Into the Reality. *American Journal of Evaluation*. **28**(1), 60–75.

- Carman, J.G. (2009) Nonprofits, Funders, and Evaluation: Accountability in Action. *The American Review of Public Administration*. **39**(4), 374–390.
- Carman, J.G., Fredericks, K.A. (2008) Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation*. **2008**(119), 51–71.
- Carman, J.G., Fredericks, K.A. (2010) Evaluation Capacity and Nonprofit Organizations: Is the Glass Half-Empty or Half-Full? *American Journal of Evaluation*. **31**(1), 84–104.
- Carter, R. (1971) Clients' Resistance to Negative Findings and the Latent Conservative Function of Evaluation Studies. *The American Sociologist*. **6**(2), 118–124.
- Chelimsky, E. (2013) Balancing Evaluation Theory and Practice in the Real World. *American Journal of Evaluation*. **34**(1), 91–98.
- Chouinard, J.A. (2013) The Practice of Evaluation in Public Sector Contexts: A Response. *American Journal of Evaluation*. **34**(2), 266–269.
- Christie, C.A. (2003) What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New directions for evaluation*. **2003**(97), 7–36.
- Christie, C.A., Fleischer, D.N. (2010) Insight Into Evaluation Practice: A Content Analysis of Designs and Methods Used in Evaluation Studies Published in North American Evaluation-Focused Journals. *American Journal of Evaluation*. **31**(3), 326–346.
- Collins, J.C. (2005) *Good to great and the social sectors: why business thinking is not the answer : a monograph to accompany Good to great : why some companies make the leap--and others don't*. J. Collins.
- Coriddi, J. (2008) Editorial - Policy & Practice - A Development Education Review. *Policy & Practice: A Development Education Review*. **7**, 1–4.
- Courtney, J. (2008) Do monitoring and evaluation tools, designed to measure the improvement in the quality of primary education, constrain or enhance educational development? *International Journal of Educational Development*. **28**(5), 546–559.
- Cousins, J.B., Earl, L.M. (1995) *Participatory evaluation in education: studies in evaluation use and organizational learning*. Falmer Press.
- Cousins, J.B., Leithwood, K.A. (1986) Current Empirical Research on Evaluation Utilization. *Review of Educational Research*. **56**(3), 331–364.
- Creswell, J.W. (2007) *Qualitative inquiry and research design: choosing among five traditions*. 2nd edition. Thousand Oaks, CA: Sage Publications.
- Cronbach, L.J. ed. (1980) *Toward reform of program evaluation*. 1st ed. San Francisco: Jossey-Bass Publishers.
- Crotty, M. (1998) *The foundations of social research: meaning and perspective in the research process*. London; Thousand Oaks, Calif.: Sage Publications.

- Dane, E., Rockmann, K.W., Pratt, M.G. (2012) When should I trust my gut? Linking domain expertise to intuitive decision-making effectiveness. *Organizational Behavior and Human Decision Processes*. **119**(2), 187–194.
- DARE Forum (2004) What is DARE. *DEEEP - Developing Europeans' Engagement for the Eradication of Global Poverty*. [online]. Available from: <http://www.deeep.org/what-is-dare.html>.
- Datta, L.-E. (2001) Coming Attractions. *American Journal of Evaluation*. **22**(3), 403–408.
- Denzin, N.K. (1989) *Interpretive biography*. Newbury Park: Sage.
- Denzin, N.K., Lincoln, Y.S. (2005) *The SAGE handbook of qualitative research*. SAGE.
- DfID (2003) *Enabling effective support: responding to the challenges of the global society : a strategy of support for the global dimension in education*. [Place of publication not identified]: DfID.
- Dickson, R. (2011) *Conduct of a meta review of programme evaluations: a case study of the SEARCH Program*. Lancaster University.
- Ding, C.S. (2009) Measurement issues in designing and implementing longitudinal evaluation studies. *Educational Assessment, Evaluation and Accountability*. **21**(2), 155–171.
- Donaldson, S.I., Gooler, L.E., Scriven, M. (2002) Strategies for Managing Evaluation Anxiety: Toward a Psychology of Program Evaluation. *American Journal of Evaluation*. **23**(3), 261–273.
- Donaldson, S.I., Scriven, M. (2003) *Evaluating Social Programs and Problems : Visions for the New Millennium*.
- Dweck, C.S. (2006) *Mindset: the new psychology of success*. 1st ed. New York: Random House.
- Eisner, E.W. (1985) *The educational imagination: on the design and evaluation of school programs*. 2nd edition. Upper Saddle River, N.J: Prentice Hall.
- Emerson, R.W., Spiller, R.E., Ferguson, A.R., Slater, J., Carr, J.F. (1971) *The collected works of Ralph Waldo Emerson*. Cambridge, Mass: Belknap Press.
- EuropeAID (2004) *Project Cycle Management Guidelines*. Brussels: European Commission, EuropeAid Cooperation Office.
- Fetterman, D. (1994) Empowerment Evaluation. . **15**(1), 1–15.
- Feuerstein, M.-T. (1988) Finding the methods to fit the people: Training for participatory evaluation. *Community development journal*. **23**(1), 16–25.
- Fitzpatrick, J.L., Worthen, B.R., Sanders, J.R. (2004) *Program evaluation: alternative approaches and practical guidelines*. Pearson/Allyn and Bacon.

- Fletcher, G., Dyson, S. (2013) Evaluation as work in progress: Stories of shared learning and development. *Evaluation*. **19**(4), 419–430.
- Flores, K.S. (2008) *Youth participatory evaluation: strategies for engaging young people*. 1st ed. San Francisco: Jossey-Bass.
- Foucault, M. (1972) *The archaeology of knowledge*. London: Tavistock Publications.
- Foucault, M. (1990) *The history of sexuality. Vol. 1: An introduction*. Reprint. London: Penguin Books.
- Freeman, M., Hall, J.N. (2012) The Complexity of Practice: Participant Observation and Values Engagement in a Responsive Evaluation of a Professional Development School Partnership. *American Journal of Evaluation*. **33**(4), 483–495.
- Friese, S. (2012) *Qualitative data analysis with ATLAS.ti*. London: SAGE.
- Fullan, M. (2011) *Change Leader*. Jossey-Bass.
- García-Iriarte, E., Suarez-Balcazar, Y., Taylor-Ritzler, T., Luna, M. (2011) A Catalyst-for-Change Approach to Evaluation Capacity Building. *American Journal of Evaluation*. **32**(2), 168–182.
- Garfinkel, H. (1984) *Studies in ethnomethodology*. Cambridge, UK: Polity Press.
- Geertz, C. (1973) *The Interpretation of cultures*. [New York (N.Y.)]: Basic Books.
- Gergen, K.J. (1999) *An invitation to social construction*. London ; Thousand Oaks, Calif: Sage.
- Geva-May, I., Thorngate, W. (2003) Reducing Anxiety and Resistance in Policy and Programme Evaluations: A Socio-Psychological Analysis. *Evaluation*. **9**(2), 205–227.
- Gherardi, S. (2009) Introduction: The Critical Power of the ‘Practice Lens’. *Management Learning*. **40**(2), 115–128.
- Giddens, A. (1979) *Central Problems in Social Theory: Action, Structure, and Contradiction in Social Analysis*. University of California Press.
- Giddens, A. (1984) *The constitution of society: outline of the theory of structuration*. Berkeley: University of California Press.
- Giele, J.Z., Elder, G.H. eds. (1998) *Methods of life course research: qualitative and quantitative approaches*. Thousand Oaks, Calif: Sage Publications.
- Glass, G.V., Ellett, F.S. (1980) Evaluation Research. *Annual Review of Psychology*. **31**(1), 211–228.
- Gong, J., Wright, D. (2007) The Context of Power: Young People as Evaluators. *American Journal of Evaluation*. **28**(3), 327–333.

- Greene, J.C. (1988) Stakeholder Participation and Utilization in Program Evaluation. *Evaluation Review*. **12**(2), 91–116.
- Greene, J.C. (1999) The Inequality of Performance Measurements. *Evaluation*. **5**(2), 160–172.
- Guba, E.G., Lincoln, Y.S. (1988) Do inquiry paradigms imply inquiry methodologies? In *Qualitative approaches to evaluation in education*. New York Praeger, pp. 89–115.
- Guba, E.G., Lincoln, Y.S. (2005) Paradigmatic controversies, contradictions and emerging confluences. In *The SAGE Handbook of Qualitative Research*. Thousand Oaks, Calif.: SAGE, pp. 191–215.
- Guba, E.G., Lynham, S., Lincoln, Y.S. (2011) Paradigmatic controversies, contradictions and Emerging confluences revisited. In N. K. Denzin & Y. S. Lincoln, eds. *The SAGE Handbook of Qualitative Research*. SAGE Publications, pp. 97–128.
- Guillemin, M., Gillam, L. (2004) Ethics, Reflexivity, and ‘Ethically Important Moments’ in Research. *Qualitative Inquiry*. **10**(2), 261–280.
- Habermas, J., Burger, T. (2008) *The structural transformation of the public sphere: an inquiry into a category of bourgeois society*. reprinted. Cambridge: Polity Press.
- Hall, G.E., Loucks, S. (1978) Teacher concerns as a basis for facilitating and personalizing staff development. *Teachers College Record*. **80**(1), 36–53.
- Hart, D., Diercks-O’Brien, G., Powell, A. (2009) Exploring Stakeholder Engagement in Impact Evaluation Planning in Educational Development Work. *Evaluation*. **15**(3), 285–306.
- Henry, G.T., Mark, M.M. (2003) Toward an agenda for research on evaluation. *New directions for evaluation*. **2003**(97), 69–80.
- Hess, M. (2006) *It’s a family affair: the media evolution of global families in a digital age*. New York: OMD.
- Hochschild, A. (2000) *The time bind: when work becomes home and home becomes work*. New York: Holt.
- Holland, J., Thirkell, A. (2009) *Measuring Change and Results in Voice and Accountability Work: Social Development Direct*. DFID Working Paper 34, UK Department for International Development, London. 220 MCNEIL AND MALENA.
- Hoole, E., Patterson, T.E. (2008) Voices from the field: Evaluation as part of a learning culture. *New Directions for Evaluation*. **2008**(119), 93–113.
- Houghton, A.-M. (2003) *The ‘initial guidance learning project’: a framework for exploring the pre-entry informal learning of part time Ph.D. learners*. Lancaster University.
- House, E.R. (2013) Evaluation’s conflicted future. In S. I. Donaldson, ed. *The Future of Evaluation in Society: A Tribute to Michael Scriven*. Evaluation and Society.

- Huber, G.P., Van de Ven, A.H. eds. (1995) *Longitudinal field research methods: studying processes of organizational change*. Thousand Oaks: Sage Publications.
- Huberman, A.M., Miles, M. (2000) Data management and analysis methods. In *Handbook of Qualitative Research*. In N.K. Denzin & Y.S. Lincoln Eds. SAGE, pp. 179–210.
- Huffman, D., Thomas, K., Lawrenz, F. (2008) A Collaborative Immersion Approach to Evaluation Capacity Building. *American Journal of Evaluation*. **29**(3), 358–368.
- Jeavons, T.H. (2005) Ethical Nonprofit management. In *The Jossey-Bass Handbook of Nonprofit Leadership and Management - David Renz, Robert D. Herman - Google Books*. San Francisco: Jossey-Bass.
- Julnes, G., Mark, M.M. (1998) Evaluation as sensemaking: Knowledge construction in a realist world. *New Directions for Evaluation*. **1998**(78), 33–52.
- Kaimacuata, A. (2010) Engaging the Disengaged through Development Education Challenges and successes. *Policy & Practice, A Development Education Review*,. **Vol. 11**(Autumn 2010), 58–68.
- Karlsson, O. (2003) Evaluation Politics in Europe: Trends and Tendencies. In O. Karlsson, ed. *Studies in Educational Policy and Educational Philosophy*. 1. Sweden: Uppsala University.
- Kearns, K.P. (1996) *Managing for accountability: preserving the public trust in public and nonprofit organizations*. Jossey-Bass Publishers.
- Kellaghan, T., Stufflebeam, D.L. (2003) *International Handbook of Educational Evaluation*. Dordrecht: Springer Netherlands.
- Kelle, U., Prein, G., Bird, K. (1995) *Computer-aided qualitative data analysis: theory, methods and practice*. London ; Thousand Oaks, Calif: Sage Publications.
- Kelly, M. (2008) *Perfectly yourself: 9 lessons for enduring happiness*. New York: Ballantine Books.
- Khoo, S. (2011) The Shifting Policy Landscape of Development Education. *Policy & Practice - A Development Education Review*. (13).
- King, K. (2007) Multilateral agencies in the construction of the global agenda on education. *Comparative Education*. **43**(3), 377–391.
- Kirkhart, K.E. (2000) Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*. **2000**(88), 5–23.
- Krause, J. (2010) *European Development Education Monitoring Report*.
- Kunreuther, F., Kim, H., Rodriguez, R. (2008) *Working Across Generations Defining the Future of Nonprofit Leadership*. Hoboken: John Wiley & Sons.
- Kusek, J.Z., Rist, R.C. (2004) *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*. World Bank Publications.

- Kushner, S. (2000) *Personalizing evaluation*. London ; Thousand Oaks, Calif: SAGE.
- Laine, M. de (2000) *Fieldwork, participation and practice: ethics and dilemmas in qualitative research*. London; Thousand Oaks, Calif.: SAGE.
- Langer, E.J. (1997) *The power of mindful learning*. Cambridge, Mass: Lifelong Books/Da Capo Press.
- Latour, B. (2010) *Nous n'avons jamais été modernes: essai d'anthropologie symétrique*. Nachdr. Paris: Editions La Découverte [u.a.].
- Lave, J., Wenger, E. (1991) *Situated learning: legitimate peripheral participation*. Cambridge [England] ; New York: Cambridge University Press.
- Law, J. (2004) *After method: mess in social science research*. London ; New York: Routledge.
- Lee, C., Nowell, B. (2015) A Framework for Assessing the Performance of Nonprofit Organizations. *American Journal of Evaluation*. **36**(3), 299–319.
- LeRoux, K., Wright, N.S. (2010) Does Performance Measurement Improve Strategic Decision Making? Findings From a National Survey of Nonprofit Social Service Agencies. *Nonprofit and Voluntary Sector Quarterly*. **39**(4), 571–587.
- Lévi-Strauss, C. (2004) *The savage mind (La Pensée Sauvage)*. Repr. New York: Oxford University Press.
- Leviton, L.C. (2015) Evaluation Practice and Theory: Up and Down the Ladder of Abstraction. *American Journal of Evaluation*.
- Leviton, L.C., Hughes, E.F.X. (1981) Research On the Utilization of Evaluations: A Review and Synthesis. *Evaluation Review*. **5**(4), 525–548.
- Liddy, M. (2010) Editorial. *Policy & Practice - A Development Education Review*. (11).
- Lincoln, Y.S., Guba, E.G. (1985) *Naturalistic inquiry*. Beverly Hills, Calif.: Sage Publications.
- Love, A. (2004) Implementation Evaluation. In *Handbook of practical program evaluation - Joseph S. Wholey, Harry P. Hatry, Kathryn E. Newcomer - Google Books*.
- MacDonald, B. (1978) Evaluation and the control of education. In D. Tawney, ed. *Curriculum evaluation today: trends and implications*. Schools Council research studies. London u.a: Macmillan, pp. 125–136.
- MacDonald, M.E. (2011) *Charitable Enterprise: The Adoption of Social Enterprise Models of Practice by a UK Charity*. Lancaster University.
- Mark, M.M., Henry, G.T., Julnes, G. (2000) *Evaluation: an integrated framework for understanding, guiding, and improving policies and programs*. 1st ed. San Francisco: Jossey-Bass.

- Mark, M.M., Shotland, R.L. (1985) Stakeholder-Based Evaluation and Value Judgments. *Evaluation Review*. **9**(5), 605–626.
- Mayer, S.E. (1996) Building Community Capacity with Evaluation Activities That Empower. In *Empowerment Evaluation: Knowledge and Tools for Self-Assessment & Accountability*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc., pp. 332–376.
- Mayne, J. (1999) Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly.
- Mayne, J. (2004) Reporting on outcomes: Setting performance expectations and telling performance stories. *Canadian Journal of Program Evaluation*. **19**(1), 31–60.
- Mayne, J. (2008) Contribution analysis: An approach to exploring cause and effect.
- Mayne, J. (2010) Addressing cause and effect in simple and complex settings through contribution analysis. In *Evaluating the Complex*. Comparative Policy Evaluation. New Brunswick, NJ: Transactions Publishers.
- McCluskey, A. (2011) Evaluation as deep learning: a holistic perspective on evaluation in the Pallete project. In M. Saunders, ed. *Reconceptualising evaluative practices in HE*. Open University Press.
- McCollum, A., Bourn, D. (2001) *Measuring effectiveness in development education*. London: Development Education Association.
- Merriam, S.B. (1998) *Qualitative research and case study applications in education*. 2nd ed. San Francisco: Jossey-Bass Publishers.
- Mertens, D.M. (2001) Inclusivity and transformation: Evaluation in 2010. . **22**(3), 367–374.
- Mertens, D.M. (2005) *Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative, and mixed methods*. Thousand Oaks, Calif.: Sage Publications.
- Mintzberg, H. (1980) *The nature of managerial work*. Englewood Cliffs, N.J: Prentice-Hall.
- Moon, J.A. (2006) *Learning journals: a handbook for reflective practice and professional development*. London; New York: Routledge.
- Morra-Imas, L.G., Rist, R.C. (2009) *The road to results: designing and conducting effective development evaluations*. World Bank Publications.
- Mowles, C. (2014) Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship. *Evaluation*. **20**(2), 160–175.
- Murray, R., Newton, M. (2009) Writing retreat as structured intervention: margin or mainstream? *Higher Education Research & Development*. **28**(5), 541–553.

- Murray, V. (2010) Evaluating the effectiveness of nonprofit organisations. In D. Renz & R. D. Herman, eds. *The Jossey-Bass Handbook of Nonprofit Leadership and Management*. John Wiley and Sons.
- Myers-Walls, J.A. (2000) An Odd Couple with Promise: Researchers and Practitioners in Evaluation Settings. *Family Relations*. **49**(3), 341–347.
- Newcomer, K.E., Hatry, H.P., Wholey, J.S. (2004) Meeting the need for practical evaluation approaches. In *Handbook of practical program evaluation*. San Francisco: Jossey-Bass.
- Nightingale, E. (2013) *The strangest secret in the world*. Merchant Books.
- O'Brien, S. (2011) The Department for International Development's Approach to Development Education. *Policy & Practice - A Development Education Review*. (13).
- Owen, J.M., Rogers, P.J. (1999) *Program evaluation: forms and approaches*. St Leonards, N.S.W.: Allen & Unwin.
- Parry-Crooke, G. (2014) *Evaluation of high support community residential services for women with complex needs*. London: Department of Health.
- Parry-Crooke, G. (2014) The Value of Evaluation featuring Georgie Parry-Crooke. [online]. Available from: <https://vimeo.com/64671664> [Accessed May 28, 2016].
- Patton, M.Q. (1978) *Utilization-focused evaluation*. Beverly Hills: Sage Publications.
- Patton, M.Q. (1994) Developmental evaluation. *Evaluation practice*. **15**(3), 311–319.
- Patton, M.Q. (1997) *Utilization-focused evaluation: the new century text*. Sage Publications.
- Patton, M.Q. (2000) Overview: Language matters. *New Directions for Evaluation*. **2000**(86), 5–16.
- Patton, M.Q. (2002) *Qualitative Research and Evaluation Methods*. Thousand Oaks, CA: SAGE.
- Patton, M.Q. (2010) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. Guilford Press.
- Pawson, R., Manzano-Santaella, A. (2012) A realist diagnostic workshop. *Evaluation*. **18**(2), 176–191.
- Pawson, R., Tilley, N. (1997) *Realistic evaluation*. Sage.
- Perrin, B. (2011) What is a results/performance-based delivery system? An invited presentation to the European Parliament. *Evaluation*. **17**(4), 417–424.
- Picciotto, R. (2015) Democratic evaluation for the 21st century. *Evaluation*. **21**(2), 150–166.
- Pred, A. (1981) Social Reproduction and the Time-Geography of Everyday Life. *Geografiska Annaler. Series B, Human Geography*. **63**(1), 5.

- Pred, A. (2007) What do YOU know? The clock is ticking, the train is rolling on. *Environment and Planning D: Society and Space*. **25**(1), 5–8.
- Preskill, H., Torres, R.T. (1999) Building Capacity for Organizational Learning Through Evaluative Inquiry. *Evaluation*. **5**(1), 42–60.
- Reckwitz, A. (2002) Toward a Theory of Social Practices: A Development in Culturalist Theorizing. *European Journal of Social Theory*. **5**(2), 243–263.
- Renz, D., Herman, R.D. (2010) *The Jossey-Bass Handbook of Nonprofit Leadership and Management*. John Wiley and Sons.
- Reynolds, J., Saunders, M. (1987) Teacher responses to curriculum policy: beyond the ‘delivery’ metaphor. *Exploring teachers’ thinking*, 195–214.
- Riddell, A. (2001) *A review of 13 evaluations of REFLECT*. Cirac.
- Rip, A. (2006) A Co-Evolutionary Approach to Reflexive Governance – and its Ironies. In *Reflexive Governance for Sustainable Development*. Edward Elgar Publishing.
- Robinson, T.T., Cousins, J.B. (2004) Internal Participatory Evaluation as an Organizational Learning System: A Longitudinal Case Study.
- Rogers, P.J. (2008) Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*. **14**(1), 29–48.
- Rogers, P.J., Petrosino, A., Huebner, A.T., Hacsí, T.A. (2000) Program Theory Evaluation: Practice, Promise, and Problems.
- Røpke, I. (2009) Theories of practice — New inspiration for ecological economic studies on consumption. *Ecological Economics*. **68**(10), 2490–2497.
- Rossi, P.H., Lipsey, M.W., Freeman, H.E. (2005) *Evaluation: a systematic approach*. Thousand Oaks; London; New Delhi: Sage.
- Rubin, H.J., Rubin, I. (2012) *Qualitative interviewing: the art of hearing data*. Thousand Oaks, Calif.: SAGE.
- Salamon, L.M. (1999) *America’s nonprofit sector: a primer*. Foundation Center.
- Saldaña, J. (2002) Analyzing Change in Longitudinal Qualitative Data. *Youth Theatre Journal*. **16**(1), 1–17.
- Saldaña, J. (2009) *The coding manual for qualitative researchers*. Los Angeles, Calif: Sage.
- Sanderson, I. (2000) Evaluation in complex policy systems. *Evaluation*. **6**(4), 433–454.
- Saunders, M. (2000) Beginning an Evaluation with RUFDATA: Theorizing a Practical Approach to Evaluation Planning. *Evaluation*. **6**(1), 7–21.

- Saunders, M. (2006) The 'presence' of evaluation theory and practice in educational and social development: toward an inclusive approach. *London Review of Education*. **4**(2), 197–215.
- Saunders, M. (2011) Capturing effects of interventions, policies and programmes in the European context: A social practice perspective. *Evaluation*. **17**(1), 89–102.
- Saunders, M. (2012) The use and usability of evaluation outputs: A social practice approach. *Evaluation*. **18**(4), 421–436.
- Saunders, M., Charlier, B., Bonamy, J. (2005) Using Evaluation to Create 'Provisional Stabilities'. *Evaluation*. **11**(1), 37–54.
- Saunders, M., Trowler, P., Bamber, V. (2011) *Reconceptualising Evaluative Practices in HE*. 1st ed. Open University Press.
- Schatzki, T.R. (1996) *Social practices: a Wittgensteinian approach to human activity and the social*. New York: Cambridge University Press.
- Schatzki, T.R., Knorr-Cetina, K., Savigny, E. von eds. (2001) *The practice turn in contemporary theory*. New York: Routledge.
- Schein, E.H. (1996) Culture: The Missing Concept in Organization Studies. *Administrative Science Quarterly*. **41**(2), 229.
- Schön, D.A. (1983) *The reflective practitioner: how professionals think in action*. New York: Basic Books.
- Schutz, A. (1972) *The phenomenology of the social world*. 1st paperback ed. Evanston, Ill: Northwestern University Press.
- Schwandt, T.A. (2002) *Evaluation practice reconsidered*. New York: Peter Lang.
- Schwandt, T.A. (2005) The Centrality of Practice to Evaluation. *American Journal of Evaluation*. **26**(1), 95–105.
- Schwandt, T.A. (2015) *Evaluation foundations revisited: cultivating a life of the mind for practice*.
- Schwarz, C., Struhkamp, G. (2007) Does Evaluation Build or Destroy Trust?: Insights from Case Studies on Evaluation in Higher Education Reform. *Evaluation*. **13**(3), 323–339.
- Scriven, M. (1991) *Evaluation thesaurus*. Sage Publications.
- Scriven, M. (2012) The logic of valuing. *New Directions for Evaluation*. **2012**(133), 17–28.
- Seppanen-Jarvela, R. (2004) The Meaning Assigned to Evaluation by Project Staff: Analysis from the Project-management Perspective in the Field of Social Welfare and Healthcare in Finland. *Evaluation*. **10**(4), 430–439.
- Shadish, W.R. (1998) Evaluation theory is who we are. *American Journal of Evaluation*. **19**(1), 1–19.

Shadish, W.R., Cook, T.D., Leviton, L.C. eds. (2000) *Foundations of program evaluation: theories of practice*. Reprinted. Newbury Park: Sage Publications.

Shadish, W.R., Epstein, R. (1987) Patterns of Program Evaluation Practice Among Members of the Evaluation Research Society and Evaluation Network. *Evaluation Review*. **11**(5), 555–590.

Shaw, I., Greene, J.C., Mark, M.M. eds. (2006) *Handbook of evaluation: policies, programs and practices*. London ; Thousand Oaks, Calif: SAGE.

Shove, E. (2003) *Comfort, cleanliness and convenience the social organization of normality*. Oxford, England; New York: Berg.

Shove, E. ed. (2007) *The design of everyday life*. New York, NY: Berg.

Shove, E. ed. (2014) *Sustainable practices: social theory and climate change*. 1. issued in paperback. London: Routledge.

Shove, E., Pantzar, M., Watson, M. (2012) *The dynamics of social practice: everyday life and how it changes*. Los Angeles: SAGE.

Shove, E., Trentmann, F., Wilk, R.R. eds. (2009) *Time, consumption and everyday life: practice, materiality and culture*. Oxford ; New York: Berg.

Silverman, D. (2011) *Doing qualitative research: a practical handbook*. Los Angeles: SAGE.

Simon, D., Knie, A. (2013) Can evaluation contribute to the organizational development of academic institutions? An international comparison. *Evaluation*. **19**(4), 402–418.

Simons, H. (1987) *Getting to know schools in a democracy: the politics and process of evaluation*. London ; New York: Falmer Press.

Simons, H. (1996) The Paradox of Case Study. *Cambridge Journal of Education*. **26**(2), 225–240.

Simons, H. (2003) Evidence-based practice: panacea or over promise? *Research Papers in Education*. **18**(4), 303–311.

Simons, H. (2009) *Case study research in practice*. Los Angeles ; London: SAGE.

Simons, H. (2015) Interpret in context: Generalizing from the single case in evaluation. *Evaluation*. **21**(2), 173–188.

Simons, H., Kushner, S., Jones, K., James, D. (2003) From evidence-based practice to practice-based evidence: the idea of situated generalisation. *Research Papers in Education*. **18**(4), 347–364.

Simons, H., McCormack, B. (2007) Integrating Arts-Based Inquiry in Evaluation Methodology: Opportunities and Challenges. *Qualitative Inquiry*. **13**(2), 292–311.

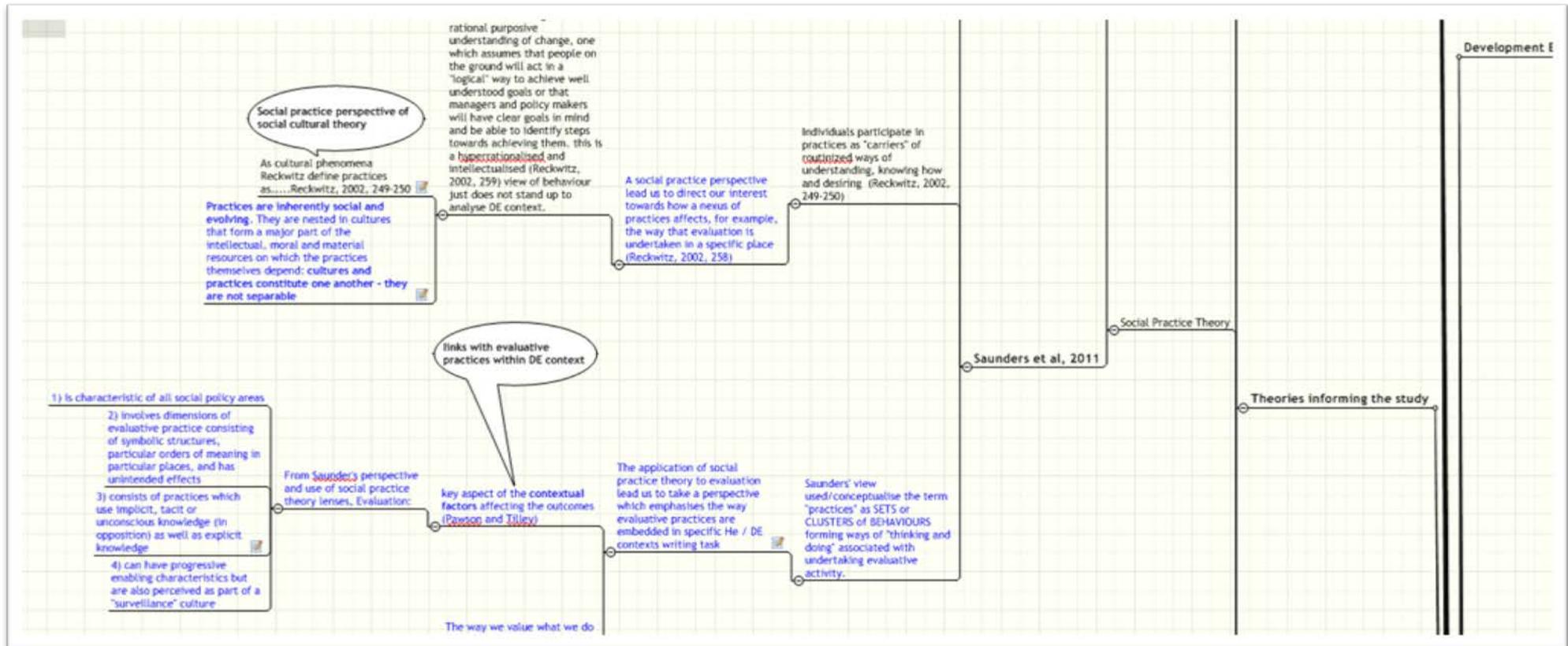
Simpson, J.A., PR, O.U., Weiner, E.S.C. (2009) *The Oxford English Dictionary*. OXFORD UK.

- Smith, N.L. (1993) Improving Evaluation Theory Through The Empirical Study of Evaluation Practice. *American Journal of Evaluation*. **14**(3), 237–242.
- Smith, N.L. (2009) Fundamental Evaluation Issues in a global society. In K. E. Ryan & J. B. Cousins, eds. *The SAGE international handbook of educational evaluation*. Los Angeles: SAGE Publications, pp. 37–52.
- Smith, N.L., Brandon, P.R. (2008) *Fundamental issues in evaluation*. New York: Guilford Press.
- Smith, S. (2011) *SME leaders' learning in networked learning: An Actor-network theory and Communities of Practice theory informed analysis*. Lancaster University.
- Southerton, D. (2003) 'Squeezing Time': Allocating Practices, Coordinating Networks and Scheduling Society. *Time & Society*. **12**(1), 5–25.
- Stake, R.E. (1983) Stakeholder influence in the evaluation of Cities-in-Schools. *New Directions for Program Evaluation*. **17**(1), 15–30.
- Stake, R.E. (1995) *The art of case study research*. Thousand Oaks: Sage Publications.
- Stake, R.E. (2004) *Standards-based & responsive evaluation*. Thousand Oaks, Calif.: Sage.
- Stake, R.E. (2005) *Multiple Case Study Analysis*. New Ed. Guilford Press.
- Stake, R.E., Kerr, D. (1995) Rene Magritte, Constructivism, and the researcher as interpreter. *Educational Theory*. **45**(1), 55–61.
- Stame, N. (2004) Theory-Based Evaluation and Types of Complexity. *Evaluation*. **10**(1), 58–76.
- Stevenson, J.F., Florin, P., Mills, D.S., Andrade, M. (2002) Building evaluation capacity in human service organizations: a case study. *Evaluation and Program Planning*. **25**(3), 233–243.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., Thomas, S. (2006) Professional Learning Communities: A Review of the Literature. *Journal of Educational Change*. **7**(4), 221–258.
- Taut, S., Brauns, D. (2003) Resistance to Evaluation: A Psychological Perspective. *Evaluation*. **9**(3), 247–264.
- Thomson, D.E. (2010) Exploring the Role of Funders' Performance Reporting Mandates in Nonprofit Performance Measurement. *Nonprofit and Voluntary Sector Quarterly*. **39**(4), 611–629.
- Trowler, P. (2013) Practice-focused ethnographies of higher education: method/ological corollaries of a social practice perspective. *European Journal of Higher Education*. **4**(1), 18–29.
- Trowler, P., Ashwin, P., Saunders, M. (2014) *The role of HEFCE in TL Enhancement*.

- Trowler, P., Saunders, M., Bamber, V. (2012) *Tribes and territories in the 21st century: rethinking the significance of disciplines in higher education*. Abingdon : Routledge.
- Trowler, P.R. (1996) *Academic responses to policy change in a single institution: a case study of attitudes and behaviour, related to the implementation of curriculum policy in an expanded higher education context during a period of resource constraint*. Lancaster University.
- Turner, J.C. (2005) Explaining the nature of power: a three-process theory. *European Journal of Social Psychology*. **35**(1), 1–22.
- Turner, S. (2001) Throwing out the tacit rule book: learning and practice. In *The Practice Turn in Social Theory*. Routledge, pp. 120–131.
- Vo, A.T. (2013) Visualizing context through theory deconstruction: A content analysis of three bodies of evaluation theory literature. *Evaluation and Program Planning*. **38**, 44–52.
- Vo, A.T., Christie, C.A. (2015) Advancing Research on Evaluation Through the Study of Context: Advancing Research on Evaluation Through the Study of Context. *New Directions for Evaluation*. **2015**(148), 43–55.
- Weiss, R.S. (1995) *Learning from strangers: the art and method of qualitative interview studies*. Simon and Schuster.
- Wenger, E. (1998) *Communities of practice: learning, meaning, and identity*. 1st ed. Cambridge: Cambridge Univ. Press.
- Wenger, E. ed. (2015) *Learning in landscapes of practice: boundaries, identity, and knowledgeability in practice-based learning*. London ; New York: Routledge.
- Wenger, E.C., Snyder, W.M. (2000) Communities of practice: the organizational frontier. *Harvard Business Review*,, 139 – 145.
- Westhorp, G. (2013) Developing complexity-consistent theory in a realist investigation. *Evaluation*. **19**(4), 364–382.
- Westley, F., Patton, M.Q., Zimmerman, B. (2006) *Getting to maybe: how the world is changed*. [Toronto]: Random House Canada.
- Wholey, J.S., Hatry, H.P., Newcomer, K.E. (2010) *Handbook of Practical Program Evaluation*. John Wiley and Sons.
- Wolcott, H.F. (1994) *Transforming qualitative data: description, analysis, and interpretation*. Thousand Oaks, Calif.: Sage Publications.
- Wolcott, H.F. (2008) *Ethnography: a way of seeing*. 2nd ed. Lanham, MD: Altamira Press.
- Worthen, B.R., Sanders, J.R., Fitzpatrick, J.L. (2004) *Program evaluation: alternative approaches and practical guidelines*. 3rd ed. Boston: Allyn and Bacon.
- Yin, R.K. (2008) *Case Study Research: Design and Methods*. 4th ed. Sage Publications, Inc.

Appendices

Appendix 1: Mind map with sample of the literature review



Appendix 2: Development Education

2a: Background information on Development Education

Development education is a commonly used term for educational work with young people, children and adults around global inequalities and injustice. Its many definitions illustrate the sometimes ad hoc nature of its growth and the diversity of the concept itself. In 1975 the United Nations defined development education as; educational work that seeks: “To enable people to participate in the development of their own community, their nation and their world as a whole. Such participation implies a critical awareness of local, national and international situations based on an understanding of the social, economic and political processes”

It is “concerned with issues of human rights, dignity, self-reliance and social justice in both developed and developing countries. It is concerned with the causes of underdevelopment and the promotion of an understanding of what is involved in development, of how different countries go about undertaking development and of the reasons for and ways of achieving a new international economic and social order.”

Both these definitions are subject to debate. However, both see development education as acquiring an understanding of ‘the causes of underdevelopment’, and its aim as taking action to create a ‘new international economic and social order’ or bring about ‘change’.

In the UK, DEA sees development education as encompassing the following principles:

- Enabling people to understand the links between their own lives and those of people throughout the world.
- Increasing understanding of the economic, social, political and environmental forces that shape our lives.
- Developing the skills, attitudes and values that enable people to work together to bring about change and take control of their own lives.
- Working towards achieving a more just and sustainable world in which power and resources are more equitably shared.

More recently DEA has refocused its work not on the advocacy of development education per se but rather ‘Global learning’ and a set of outcomes, which aim to foster: ‘critical and creative thinking; self-awareness and open-mindedness towards difference; understanding of global issues and power relationships; and optimism and action for a better world.’

This is in a large part a reflection of the **contested and nebulous nature of the term development education** but it also signifies a move away from an emphasis on delivery (traditionally via the formal education setting) to a focus on outcomes which can equally be delivered via non or informal education.

In (Adams, 2010)

Appendix 2b: Example of logic framework used by Dfid

OBJECTIVES	INDICATORS	MEANS OF VERIFICATION	ASSUMPTIONS
<p>GOAL</p> <p>The Goal is a statement of how the project or program will contribute to the solution of the problem (or problems) of the sector.</p>	<p>The indicators at Goal level describe how the overall impact of the project shall be measured. They are specific in terms of quantity, quality, and time (target group and location if relevant).</p>	<p>The means of verification are the sources of information that can be used to verify that the targets were achieved. They can include published material, visual inspection, sample surveys, etc.</p>	<p>The assumptions indicate the important events, conditions, or decisions necessary for the sustainability in the long run of the benefits generated by the project.</p>
<p>PURPOSE</p> <p>The Purpose is the direct impact to be achieved as a result of the Outputs/components/ components produced by the project. It is a hypothesis about the impact or benefit that the project attempts to achieve.</p>	<p>The indicators at the Purpose level describe how the direct impact of the project shall be measured. They should include targets reflecting the end of project status (EOPS). They are specific in terms of quantity, quality, and time (target group and location if relevant)..</p>	<p>The means of verification are the sources to which that the executor and evaluator can refer to see if the targets are being achieved. They can indicate that there is a problem and suggest the need for changes in project Outputs/components. They can include published material, visual inspection, sample surveys, etc.</p>	<p>The assumptions indicate the events, conditions, or decisions that must occur in order for the project to contribute significantly to the achievement of the Goal.</p>
<p>OUTPUTS/COMPONENTS</p> <p>These are the goods, services, and training that the project executor is required by contract to complete. They should be expressed as work completed (systems installed, people trained, etc.).</p>	<p>The indicators for Outputs/components are succinct, but clear, descriptions of each of the Outputs/components that have to be completed during execution. Each should specify quantity, quality and timing of the goods, services, etc, to be delivered.</p>	<p>This cell tells where an evaluator can find the sources of information to verify that the products/services contracted have been delivered. Sources can include site inspection, auditor's reports, etc.</p>	<p>The assumptions are the events, conditions, or decisions that have to occur in order for the Outputs/components to achieve the Purpose for which they were undertaken.</p>
<p>ACTIVITIES</p> <p>Activities are the tasks that the executor must carry out in order to produce each of the Outputs/components of the project and that denote costs. Activities are listed in chronological order for each Output.</p>	<p>This cell contains the budget for each Output/component produced by the project.</p>	<p>This cell tells where an evaluator can obtain information on whether the budget was spent as planned. It is usually the accounting records of the executing unit.</p>	<p>The assumptions are the events, conditions, or decisions that have to occur in order to complete the Outputs/components of the project.</p>

Appendix 3: Glossary

Term	Meaning in the thesis	Related debate	Informed by
Contribution	It is the value, merit or worth, as a way of finding out how a certain implementation contributes to the achievement of outcomes	Attribution as a way to support that the achievement of outcomes can be directly attributed to a specific intervention (causality) Contribution analysis builds on the notion of using a program's theory of change to infer causation. The result of a contribution analysis is not definitive proof, but rather evidence and argumentation from which it is reasonable to conclude the program has made an important contribution and why, within some level of confidence.	(Mayne, 2001, 2010; Stame, 2010; Saunders <i>et al.</i> , 2011)
Culture	Shared norms, values, and assumptions in relation to evaluation.	The internal and external environments of an organisation as "a pattern of shared basic assumptions that the group has learned as it solved problems of external adaptation and internal integration" (Schein, 1996, p. 12).	(Schein, 1996)
Competence	The skills required for participation; implying knowing and understanding, for example the skills needed to engage and participate in evaluation.	Literature refers to "capacity" in evaluation or evaluation capacity building, as having the <i>performance</i> needed to evaluate, alongside the understanding of the measurement indicators	(Carman, 2007, 2009; Shove <i>et al.</i> , 2012)
Performance	Performance based indicators, measurement approach and project cycle management	Established results-based oriented framework	(Morra-Imas and Rist, 2009; LeRoux and Wright, 2010; Thomson, 2010)
Practice	A 'practice' is "a routinized type of behavior which consists of several elements, interconnected to one other: forms of bodily activities, forms of mental activities, 'things' and their use, a background knowledge in the form of understanding, know-how, states of emotion and motivational knowledge	Conceptualisation of the elements of practice by different theorists (chapter 3)	(Reckwitz, 2002; Shove <i>et al.</i> , 2012)
Project	A project is a series of interventions aimed at bringing about clearly specified objectives within a defined time-period and budget.	There are situations in which interventions with different timescales are termed projects, and funded under specific schemes, because most external agencies have project cycle guidelines set to conceptualise these schemes	(Europe AID, 2004)
Stakeholders	People whose lives are affected by the project and people whose decisions can affect the future of the project. For the purpose of the thesis, stakeholders broadly refer to coordinators; practitioners; funders; volunteers, and trustees members, within a DE context.	Within the thesis, include practitioners as staff of the organization that practice evaluation, such as project workers, volunteers, and financial officer and youth workers.	(Gold, 1981; 1983 cited in Greene, 1988; Wenger, 1998; Saunders, 2000)

Appendix 4: Data collection

4a. Research's invitation letter

Name of Project: An exploration of the evaluation process of development education organisations: a case study

Researcher: Joana Zozimo, Phd Student, Department of Educational Research, D67, County South, Lancaster University, Tel +44 7578609287

Supervisor: Dr. Ann-Marie Houghton, Department of Educational Research, County South, Lancaster University, Tel 01524 592907

Dear participant,

20th August 2012

I would like to invite you to take part in a research study which is part of my doctoral studies in the Department of Educational Research at the University of Lancaster. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish. Ask me if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

The purpose of this study is to understand how development education organisations evaluate their interventions. My research focus is on the analysis of the Youth Project (YP) ongoing evaluation. I will describe the process in terms of monitoring current activities, outcomes measurement, evaluation planning/design, influences, use and purpose of evaluation.

Why have I been invited?

This study will focus on the ongoing evaluation process of the one of NDEC projects – the Youth Project. As a NDEC staff member I would like to invite you to take part of this study.

Do I have to take part?

The information in this letter is provided for you to decide whether you wish to participate in the present study. You should be aware that you are free to decide not to participate or to withdraw at any time without affecting your relationship with the Lancaster University research team (researcher and supervisor).

Participant's Involvement:

The research procedure is to analyse the ongoing evaluation process of the YP as a single in depth case study. At this stage I will be collecting data in your office two mornings per week and will observe all the program events and activities with regards to the evaluation process. Data will be collected throughout the research period (taken place till January 2013) through interviews, project document analysis and observations (meetings, events and activities). The interviews will be scheduled in advance and will take approximately 60 minutes. They will take place in your office or in Lancaster University meeting rooms, depending on the convenience. In the interview there will be three different moments. I will start by asking you some general questions about you and your role in the evaluation process. We will then look at the YP evaluation process and I am going to ask you to take me through it, describing it and highlighting your contribution/task/role on it and I will finish

4b. First interview protocol

Title: An exploration of evaluation in a Development Education setting: a case study

Interviewee:

Date:

Place:

Time of Interview:

Length of interview:

(Briefly describe the project) - The purpose of this study is to understand how development education organisations evaluate their interventions. My research focus is on the analysis of the youth project's (YP) ongoing evaluation as a single in-depth case study.

(Ask consent, read it and sign it)

(Inform interviewee about what will happen) - I will start by asking you some general questions about you, your professional experience and background. We will then look at the YP's evaluation trajectory and I am going to ask you to walk me through your experience of evaluation within the YP, describe it and highlight your role on it. I will finish by asking you some general questions about evaluation and the learning element involved. At the end, I might ask you some more additional questions. There is not right or wrong, only your view is valuable. I am not here to evaluate you!

A) You and your experience

1. Can you give me a brief description of your professional experience?
2. How did you start working in this sector?
3. Any important moments in your journey?

What did you do between the last job that you have mentioned and the current one?

B) Detailed description of YP evaluation trajectory

1. Consider the name of this organisation, LGEC or NDEC. Could you clarify me what
-

do you mean by a DE organisation?

2. How will you describe the DE movement? How do you differentiate DE from Global Dimension and Global Youth Work? What distinguishes these concepts?

3. Could you walk me through how the YP has been developed since the very beginning?

4. What about the evaluation planning stage? What did you do to decide which type of evaluation is better to conduct?

5. Did you have the collaboration of an external evaluator? Could you describe what their role was?

YOU

4. Could you elaborate more about your role/contribution on the YP's evaluation process? What were the last activities of the project that you had participated?

5. Could you walk me through the last YP's team meeting? What happened there? Can you describe in detail what your role in the meeting was?

6. Can you tell me how do your colleagues and yourself measure the social change in the participant's attitudes and behaviours? Tell me what did you do in order to measure the program outcomes?

(Participants were given a list of evaluation activities and were asked to tick the ones that they know/use. Look at it and ask for some examples or describe the ticked ones)

OTHERS

7. What about the other people's role in the evaluation process?

8. Going back to the last project team's meeting...can you remember what the other people's main contributions were?

FUNDERS

9. How do you describe NDEC/PDEC/ funder's relationship? Why? Do they have any influence on NDEC/PDEC? Can you give me an example?

10. Could you walk me through the last funder's visit? What happened there? Who was there?

Your understanding of evaluation / learning

1. After all you have said about the process can you tell me what your understanding of evaluation is? What about the purpose? Tell me the first words that come to mind when you think about evaluation?

2. What is in your view the purpose of evaluation?

3. What have you learn from the evaluation process, so far?

4. What have you been learning from the last two years about evaluation? Could you give me an example, please?

5. Imagine that you are recruiting a new staff member to work in the evaluation process. What will you say about it? How will you induct him/her into the task?

6. Who should I talk to in order to learn more?

7. Would you like to add any other thoughts, ideas, anything else?

(Thanks the individual for participating in this interview. Assure him or her of the confidentiality of responses and potential future interviews)

I might ask you for a second interview more focused

4c. Second interview protocol

This is supposed to be half an hour power chat....as others that we already had...but to be recorded. I found your interview very interesting and would like to come back to specific points and ask you to clarify some aspects.

1) For example, during the 1st interview you said that “....” I wonder if you could clarify what you meant by....

1.1. What is the DE ultimate goal, in your view?;

1.2. Why are you doing DE activities for?

1.3. Who is setting the agenda? What are your views

2) You have mentioned a shift from DE non-formal experience to DE formal setting.... From your perspective, what were main implications of this shift for the practice of evaluation?

3) Regarding the purpose of evaluation you have said something like...*the results were on paper, so we can meet the targets for the funders to be pleased with that*....what do you mean by that?

4) *Learning comes from a much wider thing which is difficult to quantify*Could you elaborate more on that?

5) You have mentioned in different moments throughout the interview expressions such as *difficult* and *stressful times*could you elaborate more about those? What do you precisely mean by that?

5.1. *You have said that there is a perception that you know when one activity had impact for students even if it is hard to demonstrate results....you have said, you know very well that lasts in their heads (the experiences)*. How do you know? Which are the elements that give you that sense?

4d. Interview debriefing form

Fieldwork notes – PhD Thesis Joana Zozimo

Interviewee's name	
Project name	
Interview Date	
Location	
Length	
Interview history (i.e. is this 1 st , 2 nd interview)	

- 1. In general how did the interview go?**
How did it flow? How much of the interview is 'useful' for the research project?
- 2. In brief, what were the main findings for each theme discussed in the interview?**
- 3. Were there any themes or questions which not answered/not asked/ avoided by participant? Why?**
- 4. What themes were discussed in the most depth?**
- 5. What were the most significant or interesting discussions within the interview?**
- 6. What were the least significant or interesting discussions within the interview?**
- 7. Any improvement needed on the script or fieldwork strategy?**
- 8. Final thoughts/comments**

Appendix 5: Evaluation form

Reporting activities

- Produce reports for the board of directors
- Produce reports for funders about program activities
- Produce reports for funders about financial expenditures
- Produce annual reports

Regulatory activities

- Conduct financial audits of your books
- Review program documentation (i.e., records, case notes)
- Acquire official licenses to operate programs X
- Participate in accreditation processes
- Monitoring activities
- Experience site visits by funders or regulatory agencies
- Conduct performance reviews and evaluations of staff
- Conduct firsthand observations of program activities
- Monitor program implementation

Management strategies

- Assess whether you are meeting program goals, objectives
- Establish performance targets
- Engage in formal strategic planning processes
- Use a “balanced scorecard” management system
- Evaluation and performance measurement
- Conduct formal program evaluations of your programs
- Use a performance measurement system
- Design program “logic models”

Source: Carman (2007)

Appendix 6: Reflective boxes

6a. Making sense of noticing things, thinking about things and collecting things (NCT) as an approach to analysis

As I gained awareness of what the frequent research term 'immersing in the data' meant, I realized that some of the important things happening with my analysis journey were totally unheard of. Also, I even did not know what to call them. Things, such as conduct auto coding, give 'good' names to files when preparing research materials to be uploaded and develop more descriptive or conceptual codes into analytical ones. Somehow, I had to stop and reflect. What does that mean? This meant that I had to come back to the transcription stage, make some adjustments so that the materials will be properly prepared to be analysed using ATLAS ti. One example to illustrate what I am saying is that I had to change all my file names in order to get some analytical logic later on. Surprisingly (or not) the original file names were constituted by the date, participants' initials in capital letter and transcript word written at the end.

April 2014

Excerpt from my learning journal, April 2014

6b. Researcher's learning journal

I found this analogy to be illustrative of what I experienced during the preparation of the research materials to be coded and analysed. At that stage of the analysis I was somehow lost in the unknown, exploring different paths. However I always had a sense that whatever I will do next, I will have to learn by myself in order to explain the decisions taken and the reasons why. From now on I have a feeling that maybe I will have to skip some hallways and explicitly decide not to go further deep, although it has been a temptation. An example of what I meant by 'skipping hallways' was the difficult decision to avoid exploring themes related to the affective side of the evaluation. The most difficult part was to be consistent in drawing those boundaries. Also, the challenge kept coming when I embarked into the interpretive turn of the materials. I did my best, yet I believe that feelings and emotions walk closely together, when exploring evaluative practices.

May 2014

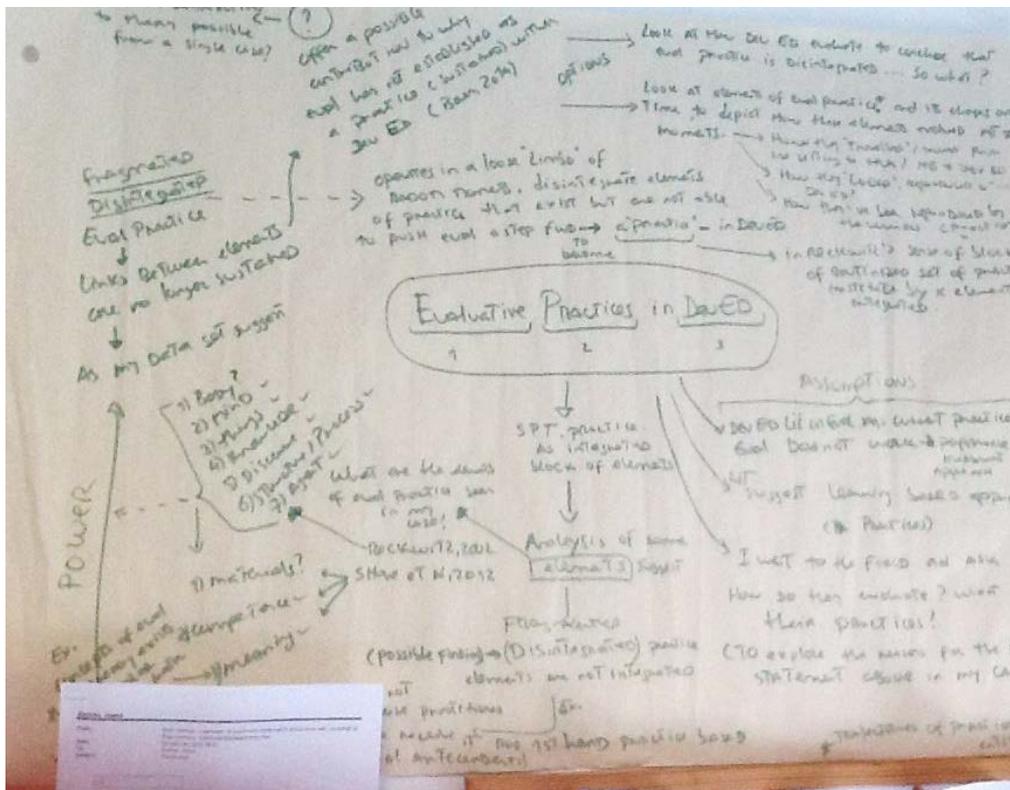
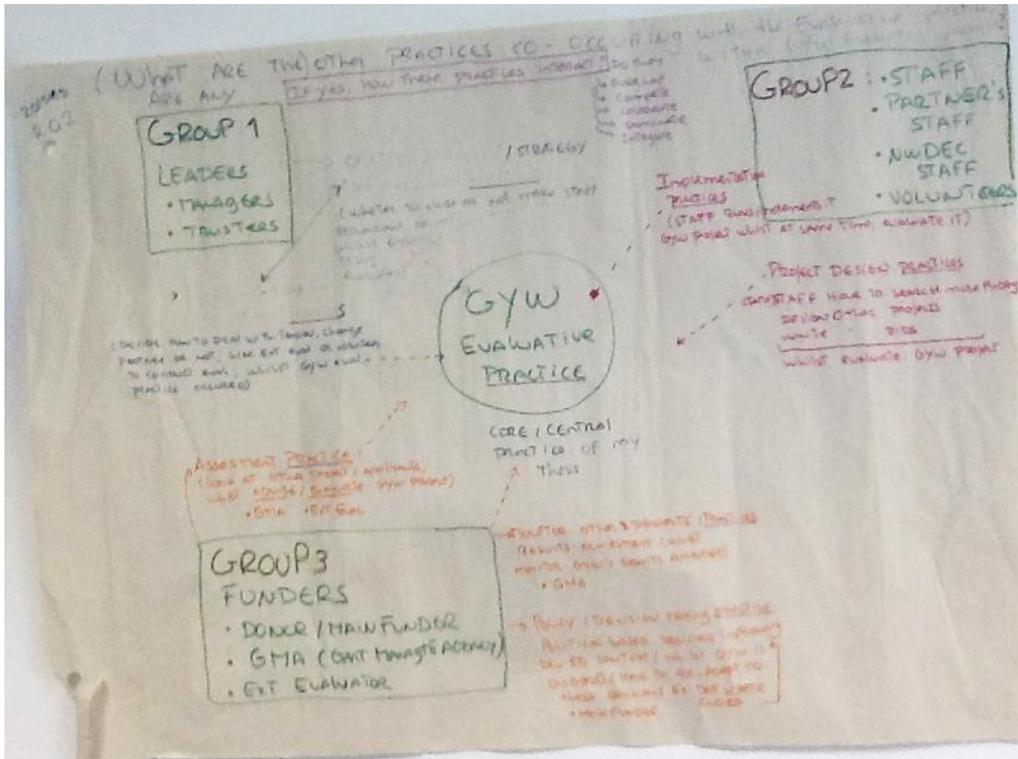
Excerpt from my learning journal, May 2014

Appendix 7: Data analysis and interpretation

7a. List of primary documents in ATLAS ti.

P-Documents	Name
P 1	Dev Ed org_manager_UA1_1st transcript_20120917.docx
P 2	Dev Ed org_manager_UA1_2nd transcript_RG_2013102.docx
P 3	Dev Ed org_Manager_UA1_Int doc_Response to GMA Y1 Report feedback_doc_2011.docx
P 4	Ded Ed org_staff_UA2_transcript_FA_20121001.docx
P 5	Dev Ed org_staff_UA2_1st transcript_HJ_20121106.docx
P 6	Dev Ed org_staff_UA2_2nd transcript_HJ_20130321.docx
P 7	Dev Ed org_staff_UA2_transcript_MS_20120925.docx
P 8	Dev Ed org_staff_UA2_1st transcript_TD_20120911.doc
P 9	Dev Ed org_staff_UA2_2nd transcript_TD_20130313.docx
P 10	Dev Ed org_staff_UA2_FN_20120515.docx
P 11	Dev Ed org_staff_UA2_FN_20120517.docx
P 12	Dev Ed org_staff_UA2_FN_20120607.docx
P 13	Dev Ed org_staff_UA2_FN_20120613.docx
P 14	Dev Ed org_staff_UA2_FN_20120620.docx
P 15	Dev Ed org_staff_UA2_FN_20120627.docx
P 16	Dev Ed org_staff_UA2_FN_20120830.docx
P 17	Dev Ed org_staff_UA2_FN_20120911.docx
P 18	Dev Ed org_staff_UA2_FN_20121015.docx
P 19	Dev Ed org_staff_UA2_FN_20130115.docx
P 20	Dev Ed org_staff_UA2_FN_20130117.docx
P 21	Dev Ed org_staff_UA2_FN_20130206.docx
P 22	Dev Ed org_staff_UA2_FN_20130214.docx
P 23	Dev Ed org_staff_UA2_FN_20130228.docx
P 24	Dev Ed org_staff_UA2_FN_20130321.docx
P 25	Dev Ed org_staff_UA2_Int doc_Eval form_2012.docx
P 26	Dev Ed org_staff_UA2_Int doc_Eval form_Outcome 3_2012.docx
P 27	Dev Ed org_Int doc_DAF small presentation draft 2010714.pdf
P 28	Dev Ed org_staff_UA2_Int doc_Eval data quotes_2012.docx
P 29	Dev Ed org_staff_UA2_Int doc_Eval tools_reporting plan_2011.docx
P 30	Dev Ed org_staff_UA2_Int doc_M&E project team meeting minutes_201104.docx
P 31	Dev Ed org_staff_UA2_Int doc_M&E notes_2011.doc
P 32	Dev Ed org_staff_UA2_Int doc_Mission statement.doc
P 33	Dev Ed org_staff_UA2_Int Strategical doc_Project Prioritising matrix_20121029.doc
P 34	Dev Ed org_staff_donor GMA_Y2 report_Doc_201106.docx
P 35	Dev Ed org_staff_UA2_Int doc_M&E plan_doc_201111.doc
P 36	Dev Ed org_staff_UA2_donor Interim progress report_Doc_201201.docx
P 37	Dev Ed org_staff_UA2_Int doc_donor_GMA visit report_doc_201201.doc
P 38	Dev Ed org_Partner Manager_UA3_transcript_B_20121016.docx
P 39	Dev Ed org_Partner Staff_UA3_transcript_WS_20121016.docx
P 40	Dev Ed org_partner_UA3_Y3 M&E doc_2013.docx
P 41	Dev Ed org_staff_UA2_donor Y3 final Proj Completion report_doc_2013.docx
P 42	Dev Ed org_partner_UA3_dissemination event report_doc.docx
P 43	Dev Ed org_Volunteer_UA4_transcript_VT_20120920.docx
P 44	Dev Ed org_Trustees_UA5_Strategical meeting_FN_20121117.docx
P 45	Donor_Ext Eval_UA6_1st transcript_HB_20121018.docx
P 46	Donor_Ext Eval_UA6_2nd transcript_HB_20121029.docx.doc
P 47	Donor_Ext Eval_UA6_NWdonor Edu network_Final Eval Report_doc_2010.pdf
P 48	Donor_HQ focal point_UA7_transcript WK_updated_20121213
P 49	Donor_GMA HQ focal point_UA9_transcript_PC updated_20121129
P 50	Donor_GMA ext advisor_UA6_enablingeffectivesupport report 2004.pdf
P 51	Donor_UA6_think tank research paper_Impact_of_global_learning_2010.pdf
P 52	Dev Ed org_staff_Int doc_Project prioritising matrix_follow up_20121029.doc
P 53	Dev Ed org_donor_External evaluation Impact report 2013.docx
P 54	GMA_UA7_10-11-0062 DAF Start up Letter 2010.pdf
P 55	Dev Ed org_partner_donor_DAF full proposal form 2009.doc
P 56	Donor_UA5_10-11 PL 0062 DFID provisional decision letter 2010.pdf
P 57	GMA_UA7_DAF Financial Management Guidelines V2.0 2010.pdf

7b. Refinement of findings using brain-assisted methods



7c. Summary of preliminary findings shared with participants

Name of Project: An exploratory case study of evaluation practice over time in a development education setting

Researcher: Joana Zozimo, Phd Student, Department of Educational Research, D67, County South, Lancaster University, Tel +44 7578609287

Supervisor: Dr. Ann-Marie Houghton, Department of Educational Research, County South, Lancaster University, Tel 01524 592907

Dear participant,

17th September 2015

I am writing to you because you took part in my research during the data collection period, in 2012 and 2013. As I am approaching a final stage of this research study, which is part of my doctoral studies in the Department of Educational Research at the University of Lancaster, I would like to invite you to read through the main findings that emerged from the study.

Please take time to reflect the following information carefully. Ask me if there is anything that is not clear or if you would like to add more information.

What was the purpose of the study?

The purpose of the study was to understand how development education organisations evaluate their interventions. My research focus was on the analysis of one of Development Awareness Fund (DAF) funded programs ongoing evaluation – in the research anonymised as the “youth project”.

Research method

This qualitative study focused on the evaluation practice of one single organisation, anonymised as NDEC. A longitudinal in-depth data collection gathered evidence from three groups of participants (staff, coordinators and funders) through semi-structured interviews, participant and non-participant observation and documentary analysis. The analysis of the data collected explored how evaluation was practiced and how it changed over time.

Research main findings

The information in this letter is provided for you to reflect on whether the findings illustrate fully or partially your perspective on the youth project’s evaluation practice. You are invited to comment or add anything you wish. However your name and the name of your organisation are not associated with the research findings in any way and your identity as a participant is known only to the researcher and her supervisor. The research data is only available to authorised personnel from the Lancaster University, in this case the researcher and her supervisor. Any published works (thesis or articles) derived from this research will never disclose your real name or the name of your organisation. The data is safely stored in

the Department of Educational Research.

The current research main findings are:

1. The circumstances under which a project is designed influences how the evaluation of that project is planned, implemented and reported – thus practised (evaluation practice).
2. Power relationships are influential to the way evaluation is perceived and practiced by DE practitioners; different power relations shape evaluation differently.
3. Evaluation is a collaborative and collective endeavour; as a multitude of concurrent practices dwell in the context of a DE project. In this context, evaluation is practised in simultaneity with other project's practices, such as management; design; implementation; and decision-making practices. Evaluation tends to be practiced collectively and not exclusively by a single practitioner.

The evaluation practice alters when new individuals (staff, coordinators or funders) join in a DE project's evaluation. It is altered because these individuals bring along other routinised daily practices - their way of doing and thinking - which interacts with the ongoing evaluation practice. As a result, various practices concur.

4. The competence of the DE practitioners to evaluate seems influential to how evaluation practice change over time, particularly over periods of unpredictable circumstances.
5. A retrospective reflection, resulting from participation in the research process, increased the awareness for the need of a practice-based evaluation (an evaluation in which the reflection on individuals' practices is central). This research highlights the benefits of reflection on evaluation practice.

How can I add my comments or feedback?

- By writing your comments in this letter and send it back by email j.zozimo@lancaster.ac.uk or post to the address below.

What will be the next steps of the research study?

I will be able to share a summary with you after the research is completed. I also aim to disseminate some recommendations for practitioners and possibly organise some workshops to share the research within a DE wider context.

Look forward to hear from you

Joana Zozimo

Phd Student, Educational Research, D67
County South, Lancaster University, LA1 4YD