BUILDING EVALUATION CAPACITY INTO A LARGE PUBLIC SECTOR EMERGENCY SERVICE: AN ACTION RESEARCH STUDY

Heidi KOLJONEN

School of Environment & Life Sciences
University of Salford, Salford, UK

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy, June 2012

Table of Contents

List of Tables

Table 1: Summary of Individual ECB approaches

Table 2: Summary of Organisational ECB Approaches

Table 3: Summary of the organisational context of ECB efforts

Table 4: Initiatives Register

Table 5: Toolkit Categories

Table 6: Research Methods

Table 7: GMFRS Barriers to Evaluation

Table 8: Focus Group Conversation Patterns

Table 9: Phases of Evaluation Activities Corresponding to Project Cycle

List of Figures

Acknowledgements

Abbreviations

BMT - Brigade Management Team

CSD - Central Service Delivery

ECB - Evaluation Capacity Building

GMFRS - Greater Manchester Fire and Rescue Service

KTP – Knowledge Transfer Partnership

PIM – Partnerships and Innovation Management

SARA - Scanning, Analysis, Response and Assessment

Abstract

This action research study investigated the feasibility of building evaluation capacity into the design process of Greater Manchester Fire and Rescue Service's (GMFRS) community initiatives. GMFRS runs over 400 community initiatives every year and an appropriate evaluation framework is needed to provide evidence of their impact and effectiveness. Information that previous evaluation processes had failed to provide.

The feasibility of establishing an evaluation framework for use by non-specialists users was explored, the development of a new evaluation system for GMFRS was examined critically, and the processes and challenges involved in embedding evaluation within a large public sector emergency service investigated.

Mixed research methods -- document analyses, qualitative interviews, observations and focus groups -- were used in the three action research cycles conducted between April 2008 and April 2010. In cycle one, existing evaluation materials and key characteristics of the initiatives were assessed. GMFRS's community initiatives were found to lack direction, and the existing evaluation tool lacked the detail required for use by personnel with no previous evaluation experience. Hence, new evaluation materials were developed and, in the second cycle, their use was observed, and interviews conducted to ascertain the barriers to evaluation practices. Lack of resources, organisational guidance, and support with evaluation activities were identified as barriers. In the third cycle focus groups were used to gain feedback on the usability of the new evaluation material and processes.

The principal conclusions of the research are that there were no theoretical models and/or guidance to assist Emergency Services to develop internal evaluation capacity. Insufficient attention has been paid to organisational support processes to nurture individuals' evaluation skills and abilities, and to assist complex organisations utilise evaluations. For evaluation to become an embedded and systematic activity, it has to be supported by a project management methodology that underpins evaluation processes.

1. INTRODUCTION TO THESIS

This chapter is divided into five sections, and begins with a statement of the purpose of the research, followed by identification of the scope and limitations of the research, a discussion about research ethics, and an overview of the thesis

1.1 Purpose of Research

The ultimate aim of the study is to develop a theoretical model for the use of evaluation capacity building (ECB) in a large public sector emergency service. The model consists of an evaluation toolkit and a series of recommendations for organisational processes to support the use of evaluation. In the thesis it is shown how to develop individual and organisational capacity to conduct and utilise evaluations, and embed internal evaluation capacity into a large public sector emergency service, something that has not been previously discussed in the ECB literature; and an argument is advanced that for evaluation to become an embedded and systematic activity, it has to be supported by a project management methodology that underpins evaluation processes. The study confirms the argument advanced by Duignan (2003), Preskill and Boyle (2008), and Taylor-Powell and Boyd (2008) that ECB has to address all the organisational aspects required to develop an evaluation system that allows individuals, with the support of the organisation, to carry out meaningful evaluations. The research responds to the need to introduce more empirically tested ECB approaches to the evaluation field (Sanders 2002, Cousins et al. 2004, Preskill and Boyle 2008), and advances the field of ECB by expanding on the models developed by McDonald et al. (2003), Boyle and Lemaire (1999) and Preskill and

Boyle (2008), by demonstrating how to embed the evaluation capacity into an organisation.

1.2 Scope and Limitations of the Study

This study originates from a Knowledge Transfer Partnership (KTP) between GMFRS and The University of Salford. The aim of the KTP project was to develop an evaluation toolkit, and to embed evaluation and feedback into GMFRS public prevention initiative design process. The KTP project was then developed into a PhD to further study organisational evaluation capacity building. A contribution is made to the ECB literature by investigating individual and organisational approaches to developing organisational ability to conduct and utilise effective and systematic evaluations. International capacity building approaches, as well as studies dealing with mainstreaming evaluation, institutionalising evaluation, evaluation culture, and empowerment evaluations are not included.

The ECB literature lacks theoretical models and/or guidance to assist emergency services and/or non-learning organisations with no previous experience of systematic evaluation practices to develop internal evaluation capacity. Hence, the research reported here focused on Greater Manchester Fire and Rescue Service (GMFRS). GMFRS runs over 400 community initiatives every year and without an appropriate evaluation framework evidence of the impact and effectiveness of the initiatives cannot be provided. If GMFRS is unable to demonstrate the impact and effectiveness of its activities, future funding and partnerships may be jeopardised; moreover the Fire Service will fail to meet internal and external quality requirements and business intelligence needs.

The research was carried out over two years, and at the end it was recognised that the organisation would benefit from the development of an evaluation culture and an initiative to mainstream evaluation. Due to the time and resource restrictions it was not possible to address these in this study, but in the discussion chapter recommendations are made for future research that addresses the limitations of this study.

1.3 Ethical Issues

Ethical approval for the research was obtained from the University's Research Ethics Panel. Taking part in all the research processes was voluntary, and the participants were able to withdraw at any time without prejudice or negative consequences. Consent was gained from everyone that chose to take part in the study, and all participants were informed of the purpose of the study and how the findings would be used. It was noted that ethical issues can get very complex in action research because the researcher is also an employee in the same organisation, and the participants are the researcher's colleagues that assist in the everyday work activities that, via the action research, also become research activities. A permission to use GMFRS data collected during the KTP and the research project in the thesis was also sought. A representative of the organisation confirmed the researcher did not need a permission for that, and that she could use the data for her PhD. The participants' confidentiality was guaranteed at all times; all the information was kept anonymous and participants are not identifiable from any published information. All the information was kept in a secure storage and only made accessible to the researcher.

1.4 Overview of the Thesis

In this thesis the various stages of developing evaluation capacity to GMFRS between April 2008 and May 2010 are examined. The literature review revealed that the current ECB models have gone as far as detailing individual and organisational processes required for successful ECB efforts, however insufficient attention has been paid to the usability of these models for non-learning organisations, and the embedding of the skills and processes into the daily activities of an organisation. This research will address these gaps, via the following research objectives:

1. to assess the process of:
    a) developing an evaluation framework and toolkit;
    b) supporting the embedding of an evaluation culture in the GMFRS against the original aims and objectives of the project, and comparing the findings to a relevant theoretical framework;

2. to investigate the challenges and benefits of non-experienced personnel using the evaluation toolkit successfully; and

3. to assess the success of a project designed to develop a self-evaluation framework and toolkit for use by non-specialists in the GMFRS.

The current state of knowledge, with regards to building evaluation capacity, is critically examined in Chapter 2, and conclusions drawn about gaps in the literature, which will be addressed in this research. The chapter concludes by arguing that the existing ECB

approaches are not suitable for non-learning public sector emergency services because of their limited nature: some of them only focus on developing individual evaluation capacity disregarding the organisational responsibilities, whereas others fail to discuss how to embed the ECB processes into the daily activities of the organisation.

In Chapter 3 an introduction to the context of the research is provided. The chapter begins by introducing the organisational context, and examining the drivers for GMFRS ECB effort. This is followed by an examination of the context of other ECB approaches, and a detailed description of the rationale for this action research study.

The theoretical debates that dominate the world of social sciences are critically examined in Chapter 4. The chapter concludes by arguing that the positivistic/constructivist dichotomy is too restrictive for applied research, hence the research utilised a pragmatic research paradigm, the methodology of action research, and mixed research methods. In Chapter 5, an examination of the research methods used in the study: document analysis, focus groups, interviews, and observations, is provided, as well as a justification for the reflective model used in the action research cycles, and a discussion about the data analysis technique.

In Chapter 6, details are given of the research activities carried out, and the findings made from the first cycle of the action research study. The aim of Cycle 1 is to understand why the existing evaluation framework is not being utilised to its full potential, and to analyse the operating environment in which the new evaluation framework will be embedded. The chapter concludes by arguing that a) the GMFRS initiatives lack direction, and b) the existing evaluation tool is not detailed enough for the use of personnel with no

previous evaluation experience. Recommendations are made for the following future actions: improve organisational understanding of evaluation, embed a project planning model which can underpin evaluation activities, develop new evaluation materials, and gain a deeper understating of barriers to evaluation.

In Chapter 7, details of the second research cycle are given. The aim of Cycle 2 is to develop an evaluation toolkit, and assess its usability amongst non-experienced personnel through three trial evaluations, and to analyse the barriers within GMFRS to the internal evaluation of community initiatives. The research activities show that a) the initiatives lack direction, b) the GMFRS personnel needs more support than just the toolkit with evaluation activities, and c) GMFRS staff members found the knowledge gap, lack of resources and organisational guidance as barriers to carrying out evaluations. The findings are followed by a decision to carry out the following actions: embed a project planning model that can underpin evaluation activities and help GMFRS personnel to clarify project goals, and to develop a wider evaluation framework to address the lack of organisational support for evaluation.

The findings of the final research cycle are discussed in Chapter 8. The aims of Cycle 3 are to review the evaluation toolkit by talking to users, and to assess the outcomes of recommendations for a wider evaluation framework and improved initiatives management system. The cycle focuses on providing evidence of how the GMFRS evaluation practices are transformed during this project, by reviewing the evaluation framework and toolkit by talking to users and stakeholders, assessing the outcomes of the attempt to improve initiative management and recording keeping system, and by

examining the organisational learning from the process of developing recommendations for a wider evaluation framework.

In Chapter 9, the results of this action research study are discussed with reference to the findings of the literature review. The chapter provides a new theoretical ECB model, concludes by arguing that that the current ECB approaches describe the functions needed for organisational evaluation capacity but do not pay attention to embedding and sustaining the evaluation skills and processes to ensure they are systematically used. Chapter 10 concludes this action research study by summarising the content of the thesis, and discussing the implications of this research to other non-learning public sector emergency services.

2. REVIEW OF CURRENT KNOWLEDGE

Evaluation Capacity Building (ECB), a sub discipline of evaluation, refers to the activity of developing processes and skills for organisations to evaluate their practices and processes. In this chapter the current state of knowledge of the subject is reviewed, and conclusions drawn about gaps in the literature that require further research. The first part of the chapter introduces the broader discipline of evaluation, and how it has developed since the 1960s. The developments in the field have required closer collaborations between evaluators and organisations which has formed the basis of ECB. In the second part, the meanings given to the term ECB are examined, and in the third section the current ECB models are critically investigated. The fourth section concludes this chapter by summarising its content.

2.1 Introduction to Evaluation

Evaluation originates from the demand for accountability and effectiveness (Scriven 1967, Alkin 1972, Rutman and Mowbray 1983, Patton 1997), and it has been a prominent feature of public sector governance since the 1960s (Patton 1997, Caracelli 2000, McNamara et al. 2009, Vedung 2010). Evaluation has no uniform and consistently applied definition (Rutman and Mowbray 1983, Lincoln 1989), and the term is challenged constantly. Evaluation is often linked to research but differs from it noticeably (Robson 2000, Boulmetis and Dutwin 2005). Evaluation assesses the value of something, whereas research describes, explains and understands (Robson 2000). Evaluators use the same methods as researchers but for different purposes. Research is concerned with understanding subjects and their relationships with what is already known in the literature,

whereas evaluation is interested in improving effectiveness of a particular project (Posavac and Carey 1985, Boulmetis and Dutwin 2005). Definitions of the term are context specific and constructed to suit the purpose, and it could be argued that the term has as many definitions as authors. Patton (2001) claims that this should not be seen as a problem, as the strength of the field is its diversity. Evaluation is impure, complex, and socially constructed (Grudens-Schuck 2003:24), and even if an agreement about the definition was achieved, there are many different ways to conduct evaluations (Patton 2001). A list of commonly used definitions of evaluation illustrates the points made in the preceding paragraph. These definitions were chosen because they represent a good variety of definitions from well-established authors; some focus on the activities of evaluation whereas others on the purpose of evaluation.

Patton (1997) and Phillips et al. (1994) offer all-inclusive definitions which detail the various activities of evaluation from the beginning till the end:

1. "Programme evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programmes to make judgements about the programme, improve programme effectiveness, and/or inform decisions about future programming." (Patton 1997:23)

2. "Evaluation is concerned with judging merit against some yardsticks. It involves the collection, analysis and interpretation of data bearing on the achievement of an organisation's goals and programme objectives." (Phillips et al. 1994:1)

Whereas Rossi and Freeman (1993) focus on the ultimate goal of evaluation:

3. "Evaluation looks at program and project objectives and asks whether they have been achieved, judges the worth of ongoing programmes, decided upon the usefulness of new programmes/projects." (Rossi and Freeman 1993:3)

Boulmetis and Dutwin (2005), similarly to Patton, focus on both the activity and outcome of evaluation:

4. "Evaluation is the systematic process of collecting data that help identify the strengths and weaknesses of a programme/project." (Boulmetis and Dutwin 2005:3)

Robson (2000) and Scriven (1991) on the other hand purely focus on the purpose of evaluation:

5. "To evaluate is to assess the worth and value of something." (Robson 2000:3)

6. "Evaluation is judging the worth or merit of something or the product of the process." (Scriven 1991:139)

As noted above, some of the definitions (1, 3 and 6) include more detail about the activities of evaluation: "collection of information", and "looking at objectives", whereas others (1, 2, 4, 5 and 6) concentrate more on the purpose of evaluation "judging the worth, "improve programmes", "decide upon usefulness", and "making decisions about future".

Boulmetis and Dutwin (2005) argue that all evaluations, no matter how defined, have one factor in common: data – which can be in the form of simple records or complex test scores. Authors also agree on the benefits of evaluation. Posavac and Carey (1985), Robson (2000), Boulmetis and Dutwin (2005), Davidson (2005), and Patton (1997) list them as: increased knowledge, improved decision making, and evidence of changes/success/impact/ efficiency/effectives/overall quality/areas for improvement. The main purpose of evaluation is to improve programme delivery, and make it more responsive to client needs (Rutman and Mowbray 1983). Evaluations determine if projects are needed and if they meet the needs of the service users (Posavac and Carey 1985). Though it has to be mentioned that evaluations are not only concerned with projects, they are also widely used to assess processes, programmes, policies, strategies, systems, services, and personnel (Davidson 2005). The fact that some of the definitions pay more attention to the actual research activity/data collection, and some to assessment of worth, defines the field. The literature is divided into approaches that mainly concentrate on the different evaluation methodologies (Scriven 1991, Rossi and Freeman 1993, Pawson and Tilley 1997, Davidson 2005), and to efforts that are primarily concerned with the utilisation of evaluations (Phillips et al. 1994, Patton 1997, Alkin 2004, Preskill (2004), Boulmetis and Dutwin 2005). Alkin (2004) argues that this is due to the underpinning reasons for the evaluation practice. Evaluations originate either from the interest for accountability or systematic social inquiry. The former is the broadest way of programme evaluation and designed to improve programmes and society. The latter is more concerned with employing a systematic and justifiable set of methods for measuring accountability (Alkin 2004).

The field has undergone considerable developments, and a shift from knowledge driven to user led approaches; the early stages were dominated by science lead evaluations, but were later superseded by more user focused methods to conduct evaluations that could be more easily used in actual decision making (Robson 2000). Vedung (2010) provides a very helpful description of the developments in the field. He refers to these as the four waves of evaluation: the first one being the scientific wave of the 1960s, in which the aim of evaluation was to make public policy and public administration more rational and grounded in facts. Evaluation was synonymous with positivistic research, distance and objectivity. The mid-1970s witnessed the 'dialogue orientated wave' when confidence in the scientific methods in social research faded, and experiments were changed to dialogues with stakeholders. During the 'neoliberal wave' of 1980s, the New Public Management, a results based management methodology was introduced and evaluation was utilised as an accountability tool. Evaluation was seen as a fundamental part of representative democracy, and the focus of evaluations was on customer satisfaction (Vedung 2010). In the 1990s evaluation was incorporated into wider management doctrines, and underwent an 'evidence wave', a return to the 1960s experimentation; the focus sifted again to making practical decisions based on research studies which adhered to strict quality criteria (Vedung 2010).

These developments have formed the basis of ECB which entails building an organisation's own capacity to conduct effective evaluations. Traditionally organisations have contracted out their evaluation needs, however now, as a way of managing limited resources, many organisation choose to build their internal evaluation capacity instead (Stevenson et al. 2002, Taylor-Powell and Boyd 2008). Stevenson et al. (2002) and McDonald et al. (2003) have noted that in most service delivery situations the programmes

are relatively small and local, and not suitable to extensive external evaluations; hence it is now quite common for funding bodies to demand that organisations develop their own internal evaluation capacity (Stevenson et al. 2002). Naccarella et al. (2007) also argue that external evaluators do not always have a full understanding of the programmes, hence fail to adequately capture what is going on. The drivers for ECB can be categorised into external and the internal demands. External demands relate to the organisations' needs to respond to the accountability demands of funders (Stevenson et al. 2002). Both public and non-profit organisations receiving external funding face the need to provide evidence of results and the impact of the money they spend (Stevenson et al. 2002, Taylor-Powell and Boyd 2008). Internationally evaluation capacity building activities have mushroomed as demands have increased for government funded programmes to demonstrate that they are effective and efficient (Naccarella et al. 2007). There are increasing expectations from funders and communities that organisations demonstrate effectiveness, value for money, and appropriateness and accessibility, both for the purposes of improving services and ensuring they are delivered optimally (Naccarella et al. 2007). The motivation to engage in ECB can also originate from internal needs and conditions (Preskill and Boyle 2008, Taylor-Powell and Boyd 2008). Organisations see evaluation as a core function in the process of becoming a learning organisation that develops the intellectual capital of staff and promotes critical inquiry (Taylor-Powell and Boyd 2008). Preskill and Boyle (2008) also list the following conditions as internal driving factors for ECB efforts:

- changes in the organisation (e.g., reorganisation, new leadership or vision);
- a mandate from leadership to increase the learning function of evaluation;
- a perceived lack of internal evaluation knowledge and skills;
- a desire to seek new or increased funding;

- a perceived shortage of evaluators with expertise and background in specific programme content or specialised organisations; and

- a desire to use evaluation to make programme improvements.

This section has reviewed the discipline of evaluation and demonstrated how the discipline has undergone a shift from knowledge driven to user led approaches to evaluation − developments that have led organisations to develop their own evaluation capacity. The next section will examine the concept of ECB in more detail.

2.2 Understanding Evaluation Capacity Building

The augmented interest in evaluation utilisation has created opportunities for closer collaborations between evaluators and organisations (Torres and Preskill 1999, Cousins et al. 2004). Evaluators have broadened the scope of their interactions with organisations and are now actively engaging in activities to better equip organisations for evaluation (McDonald et al. 2003, Cohen 2006). In addition organisations' continuing interest in improving performance and accountability has led to various efforts to build internal evaluation capability (McDonald et al. 2003, Wing 2004, Taylor-Powell and Boyd 2008). This section will explore the various meanings attributed to ECB. No consensus amongst academics and practitioners about the boundaries of the term has been reached; hence, it remains an elusive concept that accommodates a diverse range of activities. As described in the previous section, the literature is not settled even about the term 'evaluation', therefore this debate impacts upon the discussion around ECB. No two conceptions or definitions of ECB are the same (Baizerman et al. 2002).

There are various ways of approaching the task of defining ECB. Some leave much to interpretation and choose to leave it fairly vague: "ECB is the ability to conduct an effective evaluation; i.e., one that meets accepted standards of the discipline" (Milstein and Cotton 2000:1 in Naccarella et al. 2007) or "ECB is the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine" (Baizerman et al. 2002:1), or "Capacity building itself also takes in a large area including purchase of computer hardware and software, consulting assistance in everything from fundraising to strategic planning, coaching the executive director, board member training, and more" (Wing 2004:3). Whereas some opt for more detailed and inclusive notion: "[ECB] is strengthening and sustaining an organisation's capacity to (1) design, implement, and manage effective evaluation projects; (2) access, build, and use evaluative knowledge and skills; (3) cultivate a spirit of continuous organisational learning, improvement, and accountability; and (4) create awareness and support for program evaluation and self-evaluation as a performance improvement strategy (King and Volkov 2005:56 in Taylor-Powell and Boyd 2008), or "[ECB is] a context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its appropriate uses are ordinary and on-going practices within and/or between one or more organizations/programs/sites" (Stockdill et al. 2002:8). Preskill and Boyle (2008:44) have reviewed various definitions and argue that none of them truly reflect the nature of ECB and offer the following:

> "ECB involves the design and implementation of teaching and learning strategies to help individuals, groups, and organizations, learn about what constitutes effective, useful, and professional evaluation practice. The ultimate goal of ECB is sustainable evaluation practice—where members continuously ask questions that matter, collect, analyse, and interpret data, and use evaluation findings for decision-making and action. For evaluation

practice to be sustained, participants must be provided with leadership support, incentives, resources, and opportunities to transfer their learning about evaluation to their everyday work. Sustainable evaluation practice also requires the development of systems, processes, policies, and plans that help embed evaluation work into the way the organization accomplishes its mission and strategic goals."

The way Preskill and Boyle (2008) have examined the various definitions and then developed their own, characterises the ECB field. ECB definitions are (re)constructed to suit the situations at the time and reflect the needs of organisation, the evaluators and ECB developers. The key themes in the ECB definitions focus on developing evaluation skills, making institutional arrangements to support evaluation activities and the use of evaluations. The following key concepts are present in the various definitions, which could be seen as complementary rather than contradictory:

- ECB is two-fold: it enhances both 'ability to do' (potential) and actual 'doing' (practice);

- ECB is aimed at developing evaluation demand and supply;

- ECB is aimed at increasing the use of evaluation and its results;

- ECB requires development and implementation of evaluation systems;

- ECB requires institutionalization of evaluation;

- ECB could be and should be implemented at various levels: individual, group, organization, sector, national; and

- ECB is linked to creation and development of professional evaluation organisations (Kuzmin 2009:3).

In trying to define or understand the concept, it is essential to note that ECB consists of three equally important and interlinked terms: evaluation, capacity, and building. Definitions should not only make reference to equipping organisations to routinely conduct evaluations, but should also stress the varied ways of utilising evaluation findings (Patton 1997, Naccarella et al. 2007). ECB may or may not be part of "doing an evaluation, but also involve developing general awareness, skills, resources, and infrastructures to support evaluation, that is, the organizational processes that embed evaluative inquiry into the organization" (Taylor-Powell and Boyd 2008:56). Vague definitions, such as that offered by Milstein and Cotton (2001), that purely focus on the evaluation factors, fail to make this distinction and concentrate only on the knowledge and skills aspect of conducting evaluations.

As the term is constructed afresh every time it is used by authors and practitioners, and as no boundaries for ECB activities have been developed, some authors have come to view it synonymously to other evaluation related concepts: evaluation, mainstreaming evaluation, institutionalising evaluation, empowerment evaluation, and capacity building. To make the interrelated concepts more discernible they should also be examined, however ultimately all of the concepts are very similar because they are linked to the same activity: evaluation. Evaluation capacity refers to human capital skills and financial and material resources, whereas evaluation practice refers to the research design, execution of evaluation activity and the implementation of results (Duignan 2003). Mainstreaming evaluation is the process of moving evaluation to the forefront of organisational thinking and behaviour (Sanders 2002). ECB is limited to activities that improve the quality of evaluation and its use in organisations, and it may, or may not, address the goal of mainstreaming evaluation. Providing resources for evaluation is not enough to introduce

and maintain the culture of evaluation in the organisation (Sanders 2002, Grudens-Schuck 2003). Institutionalisation of evaluation means the establishment of rules, procedures and organisational arrangements by which evaluations are produced (Duignan 2003). Empowerment evaluation on the other hand is a participatory approach more closely associated with the activity of evaluation. It teaches participants to evaluate by including them in the evaluation, and increasing their capacity to plan, implement, and own programmes (Fetterman and Wandersman 2007). Capacity building, often associated with international development programmes, is a conceptual approach that concentrates on building capacity for effectively planning and implementing activities (Nu'Man et al. 2007).

Kuzmin (2009) argues that the strategy for building evaluation capacity depends on what we mean by evaluation capacity building. For the purpose of this study evaluation is defined as:

> "Evaluation is concerned with judging merit against some yardsticks. It involves the collection, analysis and interpretation of data bearing on the achievement of an organisation's goals and programme objectives." (Phillips et al. 1994:1)

Ability to evaluate in this research is defined as: a) individuals' ability to develop evaluation models, decide meaningful evaluation questions, collect useful data and make sense of the data, and b) organisational ability to utilise the data in a way that it benefits the project, all necessary stakeholders and the organisation. Evaluation capacity consists of the individuals' ability to carry out evaluations as well as the organisation's ability to provide the necessary functions required to produce evaluations. Building evaluation capacity is defined as building a system that allows an organisation to have the ability to

perform and produce functional evaluations. Meaningful and functional evaluations are well planned, and systematically produced and utilised with organisational support. Preskill and Boyle (2008) argue the core function of ECB is to build sustainable evaluation practice, hence ECB is defined as a marriage of evaluation skills, knowledge, and the organisational (human and non-human) functions to evaluate and support evaluations.

This section has examined various definitions for the term 'evaluation capacity building', and demonstrated it is a relatively abstract and novel concept. It was also shown how the term overlaps with various other evaluation related concepts, contingent on how the different terms are defined and how the ECB environment is able to accommodate them. The next section will investigate current ECB approaches, and clarify the necessary functions and abilities required to develop organisational capacity to evaluate.

2.3 Review of Existing Approaches to Building ECB in an Organisation

This section critically assesses the existing approaches to ECB. The approaches reflect the definition debate - none of the approaches, as with the definitions, has been used more than once. A new one is developed to represent each ECB effort, even though the methods the developers have employed are identical to each other. An approach in this thesis is defined as an ECB development, a study that discusses how to develop ECB into an organisation. The researcher has divided the approaches found in the literature into two categories: those that concentrate on building the evaluation skills of the individual, and those that in addition to that, recognise the need to equip the organisation for evaluations.

Most ECB authors (Stevenson et al. 2002, Lennie 2005, Monroe et al. 2005, Cohen 2006, Forss et al. 2006, Taut 2007, Adams and Dickinson 2010) concentrate on the organisations ability to produce evaluations, i.e. building the evaluation skills of the individual. Monroe et al. (2005), for example, argues that evaluators can build evaluation capacity into organisations by teaching staff to use logic models and by building their technical evaluation skills. Stevenson et al.'s (2002) approach also consists of evaluation training workshops addressing knowledge of logic models, data collection, and data analysis, as well as a separate needs assessment to identify the right type of evaluation to suit the organisation. Cohen (2006:2) similarly takes the approach that ECB is built through a shared study experience, an adult learning strategy designed to empower as well as educate. Table 1 contains details of authors' positions on individual arrangements for ECB. It can be noted from Table 1 that individual approaches consist of training: on evaluation design, logic models and data collection methods, and provision of external assistance with conducting evaluations. In some cases, training is carried out in the form of participatory evaluations, where staff are included in the evaluation process, and this way learn to master the skill of evaluation. In this thesis these approaches are called individual approaches because they equip individuals with evaluation skills, but do not discuss the organisation's responsibilities in the evaluation process.

Table 1: Summary of Individual ECB approaches

INDIVIDUAL ECB APPROACHES	
Author	**Individual Approach**
Porteous (1999)	Evaluation toolkit and an accompanying workshop.
Stevenson et al. (2002)	Training, on-site and telephone technical assistance, participatory evaluations.
Barnette and Wallis (2003)	Training – logic models, theory of change, evaluation techniques and tools.
McDonald et al. (2003)	Evaluation training and trial evaluations.
Monroe et al. (2005)	Training- evaluation tools, design, logic models and programme theory.
Arnold (2006)	Teaching to use logic models, provision of one-on-one assistance, facilitating collaborative evaluations.
Cohen (2006)	Training- evaluation tools, design and logic models.
Forss et al. (2006)	'Learning by doing' is an effective way to develop the capacity of professionals in an organisation.
Miller et al. (2006)	Staff need basic evaluation training to understand the basic processes.
Naccarella et al. (2007)	User-friendly manuals on programme evaluation, technical assistance, training workshops, interactive web-based systems to guide evaluation design, data collection, data entry and analysis.
Taut (2007)	Evaluation workshops, facilitation of self-evaluation, support materials, training of future facilitators.
Huffman et al. (2008)	Participatory evaluation, staff participate in real life evaluations and bring the knowledge back to the organisations.
Taylor-Powell and Boyd (2008)	Training, technical assistance, collaborative evaluation projects, mentoring and coaching, evaluation materials, evaluation champions.
Adams and Dickinson (2010)	Use of adult learning and experimental learning philosophies to teach theory-driven evaluation approach.

As mentioned in the preceding paragraph, the approaches listed in Table 1 represent ECB approaches that focus on equipping individual members of organisations to carry out evaluations. These approaches do not consider the organisations' responsibility in the ECB process and make the assumption that, by ensuring staff have the necessary skills to evaluate, the organisation will ultimately have the capacity to evaluate. It is

interesting to note how little the approaches vary; all of them use identical methods, but still call their ECB models unique. The common themes in the approaches are training and the provision of evaluation materials but the level and intensity of training and assistance may vary. McDonald et al. (2003) are critical of approaches that deal only with developing staff skills in data collection and analysis, and do not involve the need to tackle broader organisational change. Many efforts at building evaluation capability have focused primarily, or even exclusively, on documenting and developing the skills, tools and resources that are available to produce evaluations (McDonald et al. 2003:10). There is a need to step away from the thinking that capacity building is purely an activity that teaches people to evaluate (Huffman et al. 2008, Hay 2010), because if the organisation does not support and integrate evaluation it will not have an evaluation system, which essentially is the focus of ECB. Purely focusing on building evaluation skills can lead to a situation where an organisation may be capable of producing evaluations but unable to use them, or even worse, produce evaluations that are treated as irrelevant (Williams 2001 in McDonald et al. 2003). All the skills, knowledge, technical expertise and experience in the world will not help if the programme, community, organisation, or environment cannot sustain and nurture those skills and abilities (Williams 2001 in McDonald et al. 2003).

McDonald et al. (2003) explore the role of ECB and use the well-known epigram 'give someone a fish and they eat for a day; teach them to fish and they eat for a lifetime' (p.10). In applying the analogy to ECB, they argue that organisations need the equipment to successfully fish, an effective distribution system, people who want to eat fish, and an entire fishing system that is sustainable. They note that the efforts have to focus on working with the whole organisation, not just on developing the skills of individuals. ECB is a collaborative effort that requires the attention of everyone involved in the activity. The

aim is not just to conduct evaluations but also to commission, manage and use them (McDonald et al. 2003:10). This is an important consideration for trying to establish the variety of ECB activities. Following on from the previous section (Understanding ECB), which highlighted the work of Duignan (2003), Preskill and Boyle (2008), and Taylor-Powell and Boyd (2008) who demonstrated that ECB has to address all the organisational aspects required to develop an evaluation system that allows individuals, with the support of the organisation, to carry out meaningful evaluations. Evaluation capability should provide enduring organisational benefits, including a sustainable resource for producing evaluations as well as a system for encouraging and using evaluation. Table 2 summarises the current organisational approaches to ECB.

Table 2: Summary of Organisational ECB Approaches

ORGANISATIONAL ECB APPROACHES	
Author	**Organisational Responsibilities**
Minnett (1999)	Trained internal evaluator conducts evaluations and acts as an evaluation mentor.
Barnette and Wallis (2003)	Provision of external evaluators that work closely with programme staff.
McDonald et al. (2003)	Make evaluation mandatory for all projects.
Arnold (2006)	Recruitment of full time evaluation assistance, provision of resources (time, money, software), personal development opportunities and support for evaluation.
Miller et al. (2006)	Organisation "insources" evaluations to external evaluators that conduct the evaluations.
Huffman et al. (2008)	Opportunities for participatory evaluations
Taylor-Powell and Boyd (2008)	Leadership, demand, incentives, structures, policies and procedures, organisational assets, financing, technology, time.
Adams and Dickinson (2010)	Systems change is needed to introduce evaluative practices.

As it can be noted from Table 2, fewer authors have recognised the need for organisational responsibilities together with ECB systems that focus on wider evaluation frameworks consisting of organisational support and strategic management. The commonality between these organisational approaches is that they all recognise the importance of involving the organisation in the ECB process but the way they aim to support staff members in evaluation activities varies greatly. The approaches will be examined more closely in the subsequent paragraphs.

The focus of ECB should be on developing processes and practices that make evaluation part of the everyday work of an organisation (Huffman et al. 2008:359). Kuzmin (2009:4) argues a comprehensive ECB strategy should include:

1. Nurturing the need for evaluation services and specialists;
2. Establishing stable career opportunities in evaluation;
3. Creating and maintaining a body of knowledge and set of skills unique to evaluation;
4. Developing educational programs and other professional development opportunities for evaluators;
5. Institutionalising evaluation; and
6. Building professional evaluation associations.

The strategy is a good starting point because it recognises that an organisation has to support evaluation practices, and have a leading role in the ECB process. However, without any empirical evidence, the approach takes a very theoretical position, and does not consider what the different stages should entail, and how they should be developed in

organisational settings. Taylor-Powell and Boyd (2008:63) have similarly developed an ECB framework that does not rely exclusively on training and building knowledge and skills of individuals, but also requires the following stages to function successfully:

1. Communication structures that facilitate horizontal and vertical information flows across the entire organisation;
2. A team program structure that facilitates collective action, collaborative inquiry, group problem solving and synthesis. Additional peer-support and learning structures, such as program area liaison structures, evaluation advisory groups, and mentoring structures, which can build on existing mechanisms to facilitate ECB;
3. A data management system to facilitate creation, management, and use of data, and incorporates question banks for customised data collection, Web based data processing, templates for using and communicating data, and processes for monitoring data quality and sharing lessons learned; and
4. Policies and procedures. A variety of explicit and implicit rules and procedures guide evaluation decisions and actions.

Whereas Kuzmin (2009) recognised the need for organisational responsibility in the ECB process, this more detailed and comprehensive approach has identified what organisational support functions are required for organisational evaluation capacity. The framework also raises an important point about the collaborative aspect of ECB efforts. Stevenson et al. (2002), Forss et al. (2006), Huffman et al. (2006), Naccarella et al. (2007), Taut (2007), and Compton (2009) have mainly concentrated on the use of external assistance, not only in building evaluation capacity but also in maintaining the ECB

structures. However there is a need to address the joint individual and organisational responsibility in order for evaluation to play an expanded role in organisations (Preskill and Boyle 2008). Boyle and Lemaire (1999:13) also take a more participatory approach to ECB and argue that the following factors are essential in creating an appropriate balance between demand and supply for evaluation:

1. Independence (grant autonomy to the evaluation function and access to data);

2. Skilled evaluators;

3. Authority to access data and personnel;

4. Location in the organisation (sufficient rank structure);

5. Evaluation agenda:

   • scope of evaluation work

   • authority

   • reporting requirements

   • evaluation relationships with managers; and

6. Creditability – use of objective evaluators.

Boyle and Lemaire (1999) argue that the key to building successful evaluation capacity in organisations is the awareness of benefits and disadvantages of evaluation in assisting programme development. This emphasises the participatory aspect of ECB and highlights the importance of having a common organisational understanding of evaluation and ECB models and systems. Whilst developing ECB systems, Duignan (2003) has also noticed that limited skills and knowledge about evaluation at all levels in organisations is common, hence he argued for the importance of demystifying evaluation and having a shared understanding of all activities. Duignan (2003) also emphasised that an organisation

has to have the ability to talk sensibly about evaluation questions, in order to ensure everyone knows how to take evaluation activities appropriate to their situations. Everyone in an organisation plays a key role in the ECB process and it is highly important that the organisation comes together in deciding their evaluation needs. Duignan's contribution to the debate is important, as unless a common definition for evaluation and shared expectations for the ECB efforts are reached, all the different levels of an organisation will not be able to engage in the process.

McDonald et al. (2003) introduced two new dimensions to the ECB discussion: the notion of testing the evaluation models; and the mandatory aspect of evaluations, by recommending the following stages for ECB (p.13-19):

- Stage 1: addressing the need for a type of evaluation of the particular intervention;

- Stage 2: experimenting with volunteer projects - The second phase, a piloting phase, involved developing evaluation strategies across a range of projects to test whether the approach taken in the first phase would hold in different fields of activities and to determine what modifications might be required; and

- Stage 3: Evaluation became mandatory for all new projects, not just those involved in the new initiative. Approval for new project proposals was only given if a credible evaluation plan was provided.

McDonald et al.'s (2003) approach suggests there is an imperative to consider how to embed the evaluation capacity into an organisation. Knowledge of evaluation does not mean the organisation will have the capacity to evaluate, similarly the existence of support

functions does not guarantee they will actually be utilised. Various aspects of programme delivery are mandatory; hence evaluation should not be seen as a separate function. If the goal of the ECB effort is to build a system that makes evaluation use a routine practice, it has to be incorporated in all of the programme functions and exist as an equal counterpart to those functions. Testing ECB and/or evaluation models with the existing programme staff adds a new aspect to the evaluation training and as ECB frameworks are socially constructed, they have to suit the organisation's operating environment and the routine practices of the programme delivery.

Finally, Preskill and Boyle (2008) argue that there have been many articles and discussions about ECB over the years, but there appear to be few comprehensive conceptual frameworks or models that could be used to (a) guide practitioners' ECB efforts, and/or (b) empirically test the effectiveness of ECB processes, activities, and outcomes (p.444). They have developed a model (Figure 1) to provide a set of guidelines for designing and implementing ECB efforts:
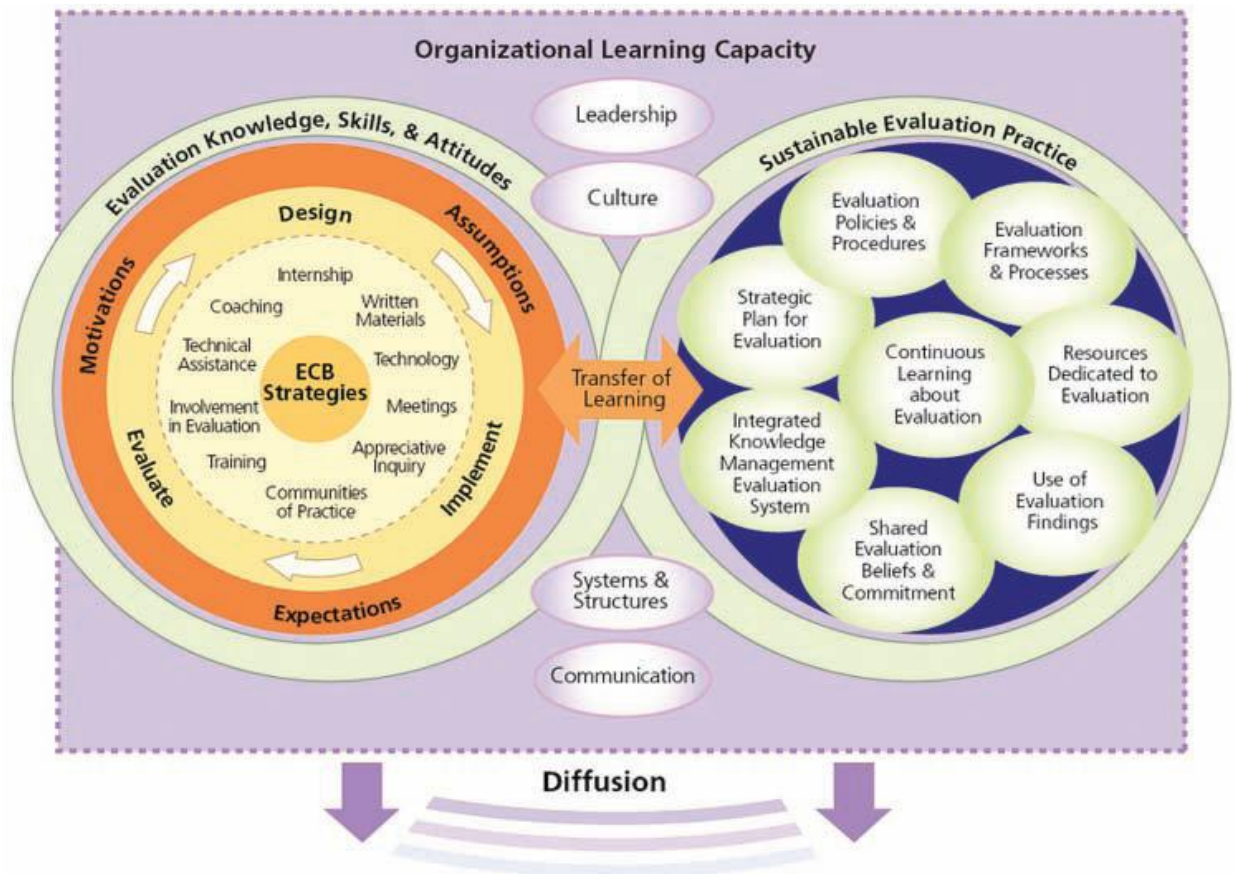
Figure 1: Preskill and Boyle (2008:446) "A Multidisciplinary Model of Evaluation Capacity Building (ECB)"

The circle on the left of the model represents the initiation, planning, design and implementation of the ECB effort, which reflects the goal of ECB being the development of evaluation knowledge, skills, and attitudes. The double-sided arrow, that connects the circles represent the transfer of learning, which refers to the application of evaluation knowledge, skills, and attitudes to the work context. The circle on the right side of the model describes the processes, practices, policies, and resources that Preskill and Boyle (2008) consider to be required for sustainable evaluation practice. The final component of the model is diffusion, which reflects the participant's ability to share their knowledge and skills with a wide range of audiences. The authors infer that as the participants share their evaluative thinking and practices, others will be inspired to learn about and engage in

evaluation practice as well (Preskill and Boyle 2008:445-446). Preskill and Boyle argue that although it may be difficult to achieve all of these in one organisation, they should be seen as goals for those who see evaluation as a means for achieving individual, group, and organisational learning (Preskill and Boyle 2008:446).

There are also models in the literature that do not fit into the researcher's categorisation of the current ECB approaches, because they do not discuss the development of organisational structures and/or skills to allow organisations to evaluate their own activities, the focus of these studies is on external assistance.. These are ECB efforts developed by Huffman et al. (2006) and Compton (2009), which use collaborations with local universities as means of developing organisational evaluation capacity. The approaches recommend organisations to make arrangements with university students to carry out evaluations for the organisation – the students get much needed work experience, and the organisation benefits from the outcomes of evaluations. While several additional studies have been identified from the literature they are not included in either of the tables due to their limited application in the present study. The work of Gibbs et al. (2002) has not been included because it discusses how funding bodies and technical assistance agencies can help build ECB in community organisations, rather than the community organisations developing their own capacity. King's (2002) ECB model has been excluded because it is developed for a district compromising of forty schools and thirteen communities, whereas this research focuses on organisational ECB; similarly Milstein et al. (2002) has also been excluded because their ECB study focuses on strengthening an existing evaluation capacity system.

Boyle and Lemaire (1999), McDonald et al. (2003) and Preskill and Boyle (2008) are the only authors that have shown some consideration to the need to embed ECB processes into the organisation in order to develop a sustainable system that is systematically utilised. However they have failed to explain how all the ECB activities will fit into the existing organisational structures, and have made no recommendations on how an organisation would implement all these additional responsibilities in their everyday operations. Most importantly, the authors do not describe whose responsibility it is to develop the necessary leadership to manage these functions. Often the external evaluator provides the expert advice and recommends organisational systems to support the use of evaluation, but it remains unknown who replaces the evaluator once he/she is gone. Cohen (2006) argued for the importance of maintaining a close relationship between the participants and the evaluator but, as with other models, does not detail how that relationship is maintained after the capacity building initiative has ended, which can have a major impact on the system's sustainability. The final part of this chapter revealed that the majority of the current ECB efforts (Stevenson et al. 2002, Monroe et al. 2005 et al., Lennie 2005, Cohen 2006, Forss et al. 2006, Taut 2007, Adams and Dickinson 2010) focus on developing individual evaluation skills and expertise. Kuzmin (2009) recognised the need to develop functions to support the conduct and use of evaluations but failed to detail the precise nature of these functions. The most comprehensive ones (McDonald et al. 2003, Boyle and Lemaire 1999, Preskill and Boyle 2008) have tried to move away from the individualistic and expert centred approaches, by recommending ways to implement the capacity into the organisations and describing structures that are essential in developing sustainable evaluations systems.

This section of the chapter has examined what is currently known about ECB – how it is developed, and what kind of individual and organisational approaches are used to develop organisations' capacity to evaluate. The individual ECB approaches focused on equipping staff members to carry out evaluations by teaching them how to conduct evaluations. Furthermore the organisational ECB approaches went somewhat further by arguing that in addition to equipping the individuals, the organisation also has to commit to the evaluation process by providing leadership, technology, communication channels to utilise evaluation findings, and policy and procedures to set the "rules and regulations" of evaluation. All the approaches examined in this thesis have produced the desired outcomes in the environments within which they were developed. As mentioned before, ECB approaches are constructed to suit the context; hence different approaches suit different situations. Currently, there is no empirical evidence about the effectiveness of these approaches, which will be discussed in more detail in the next chapter. The review of the literature on current ECB models has raised the following issues that will be investigated:

- the current ECB approaches are designed for external evaluators, not for organisations hoping to develop their own evaluation capacity, skills and expertise;

- the current ECB approaches utilise a variety of methods to build organisational evaluation capacity, and describe the functions needed for successful ECB but none of them discusses how to sustain the evaluation skills and processes, and ensure they are systematically used; and

- there is also a gap in the knowledge of how to embed evaluation capacity into an organisation that does not have the funds or capacity to accommodate the additional support functions/organisational processes identified by the authors.

2.4 Conclusion

In this chapter the field of evaluation, ECB definitions and current ECB models have been examined. The study has revealed that evaluation and ECB are fluid, contextually constructed concepts that derive from the organisational need for more effective practice and improved accountability. It has been also identified that ECB definitions should consider the individual's ability to conduct evaluations, the organisational capacity to use evaluations and the various stages of building the necessary processes that accommodate and support both individual and organisational capacity to evaluate. Building evaluation skills and knowledge potentially is an integral part of ECB efforts, however ECB goes beyond that. ECB cannot be limited to the activity of skills building, as it also deals with the organisational processes that support the knowledge and skills of those who evaluate. Hence, within this chapter it was argued that ECB is much more than building skills to evaluate. Evaluation as a standalone activity cannot function without the support of the organisation; ECB definitions have to include notions of activities that encourage organisations to build systems that assist in the conduct of useful evaluations. The current ECB models were divided into approaches that develop individual capacity to conduct evaluation, and to models that, in addition to the individual skills, equip organisations to support the conduct and utilise evaluations. Within this chapter, gaps were identified about the sustainability of the existing models, and the lack of

contextual detail. The questions raised following review of the literature will be addressed in this study.

For the purpose of this work, the researcher used Phillips et al (1994:1) definition of evaluation: "Evaluation is concerned with judging merit against some yardsticks. It involves the collection, analysis and interpretation of data bearing on the achievement of an organisation's goals and programme objectives", because it covers all aspects of evaluation activity – the purpose, the actions and intended outcomes. After an examination of the existing ECB definitions the researcher defined ECB as a marriage of evaluation skills, knowledge, and the organisational (human and non-human) functions to evaluate and to support such evaluations.

3. CONTEXTUALISING THE RESEARCH

In this chapter, an introduction to the context of the research is provided. The chapter is divided into six sections, and begins by introducing the organisation, and its role in the wider community. This is followed by a discussion about the organisational need for evaluation, and the background to the research project. In the fourth section the contexts in which the current ECB models have been developed are examined, and in the fifth aims and objectives are set for this action research study. The final section concludes the chapter.

3.1 Introduction to the Organisation

GMFRS is a large public sector emergency service that employs approximately 2,400 people, of which 70% are uniformed operational staff (GMFRS 2011). The work they carry out ranges from fighting fires and rescuing people, to promoting fire safety to the wider communities. GMFRS, the second largest fire service in the UK, consists of central management that deals with managerial and support functions, and ten Boroughs - Bolton, Bury, Manchester, Oldham, Rochdale, Salford, Stockport, Tameside, Trafford and Wigan – which mainly concentrate on the operational side of the organisation's activities. GMFRS is governed by Greater Manchester Fire and Rescue Authority, and funded by central government grants and council tax contributions (GMFRS 2011). The organisation's core purpose is to protect and improve the quality of life of the people in Greater Manchester, and this is achieved via the following aims (GMFRS 2011:4):

1. Prevention - Engage with Greater Manchester's communities to inform and educate people in how to reduce the risk of fires and other emergencies and do all we can to prevent crime and disorder;

2. Protection - Influence and regulate the built environment to protect people, property and the environment from harm;

3. Response - Plan and prepare for emergencies that may happen and make a high quality, effective and resilient response to them;

4. People - Work with people with the right skills and attitude to deliver high quality, value for money services in a positive environment for everyone;

5. Public Value - Manage risk through using resources flexibly, efficiently and effectively, continuously improving our use of public money in ways the public value; and

6. Principles - Operate in accordance with the law and our values, and ensure that safety, sustainability, partnership and inclusivity run through all we do.

The GMFRS aims range from preventing fires from occurring to protecting the public from possible incidents. However the aims have not always been as varied as they are now. Prior to 2002 the organisation's main emphasis was on fighting fires (Fire Service Act 1947), however in 2002 an independent review revealed that, in addition to institutional and management structure changes, Fire and Rescue Services' approach to fire prevention and community fire safety needed to be modernised:

"A radical programme of reform is required to change the Fire Service into a modern institution which can truly deserve the trust, confidence and respect which the public places in it". (Bain, Lyons and Young 2002:3)

"The future responsibilities of the Fire Service should include partnership and collaboration with other agencies, including the stationing of emergency vehicles and staff at existing fire stations, and acceptance of an increased role in medical assistance and use of resuscitation equipment by fire crews. The Fire Service must deliver fire safety services in partnership with community and local organisations. The move to a risk-based assessment of fire safety must be supported by the flexible deployment of resources to maximise the efficiency of the Fire Service. These changes must be supported by legislative and funding reform by central government". (Bain, Lyons and Young 2002:10)

As demonstrated in the quotes above, the Fire and Rescue Services' approach to fire prevention, and the way they interact with the wider community, needed to be modernised to include more collaborations with other organisation. The recommendations were followed up by the Fire Service Act of 2004 that reinforced the findings, and set a new agenda for Fire and Rescue Services across England which included the new responsibilities for fire safety (Fire Service Act 2004: 9):

1. A fire and rescue authority must make provision for the purpose of promoting fire safety in its area; and

2. In making provision under subsection (1) a fire and rescue authority must in particular, to the extent that it considers it reasonable to do so, make arrangements for:

    a) the provision of information, publicity and encouragement in respect of the steps to be taken to prevent fires and death or injury by fire;

    b) the giving of advice, on request, about;

    c) how to prevent fires and restrict their spread in buildings and other property; and

    d) the means of escape from buildings and other property in case of fire.

As a result of the new legislation, for the first time in the organisation's history, the focus of their key activities had shifted from response to prevention. To fully reflect the changes brought about by the modernisation of the Fire and Rescue service, GMFRS developed a vision "to make Greater Manchester a safer place by being a modern, community focused and influential Fire and Rescue Authority" (Greater Manchester Fire and Rescue Authority 2004). As part of the Fire and Rescue Service modernisation programme, GMFRS also adopted a borough based model to improve performance in preventing fires, enhance community fire safety and to collaborate more widely with local partners. The organisation was divided in ten Boroughs, to represent the ten local Boroughs of Greater Manchester, which were given a certain degree of independence from the central management, to allow them to make effective decisions about their local areas (Greater Manchester Fire and Rescue Authority 2004).

3.2 Greater Manchester Fire and Rescue Service and Community Work

All of the newly emerged Boroughs were required to enter into partnerships with other local bodies, and to develop and deliver numerous public protection initiatives and partnerships in order to meet the new external demands for making communities safer. GMFRS believed the underlying causes of fire closely correlated to those leading to crime: anti-social behaviour; poor health, and educational underachievement, and decided to make a major contribution towards partnerships that educate and inform children and young people about the dangers of fire, fire-related crime, and its consequences (KTP Grant Application and Proposal Form 2007). Therefore the organisation initiated a large number of diverse projects designed to engage with and influence the communities in order to achieve a safer society. These initiatives included age group specific interventions

designed to deter fire setting behaviours, awareness raising activities of the risks associated with car crime, schools interventions aimed at re-engaging young people with education, and initiatives engaging with young offenders aimed at modifying the behaviour of at risk young people (KTP Grant Application and Proposal Form 2007:6). The management and delivery of the newly developed community activities did not concern the whole organisation; they came under the Fire Safety department, as shown in Figure 2[1].
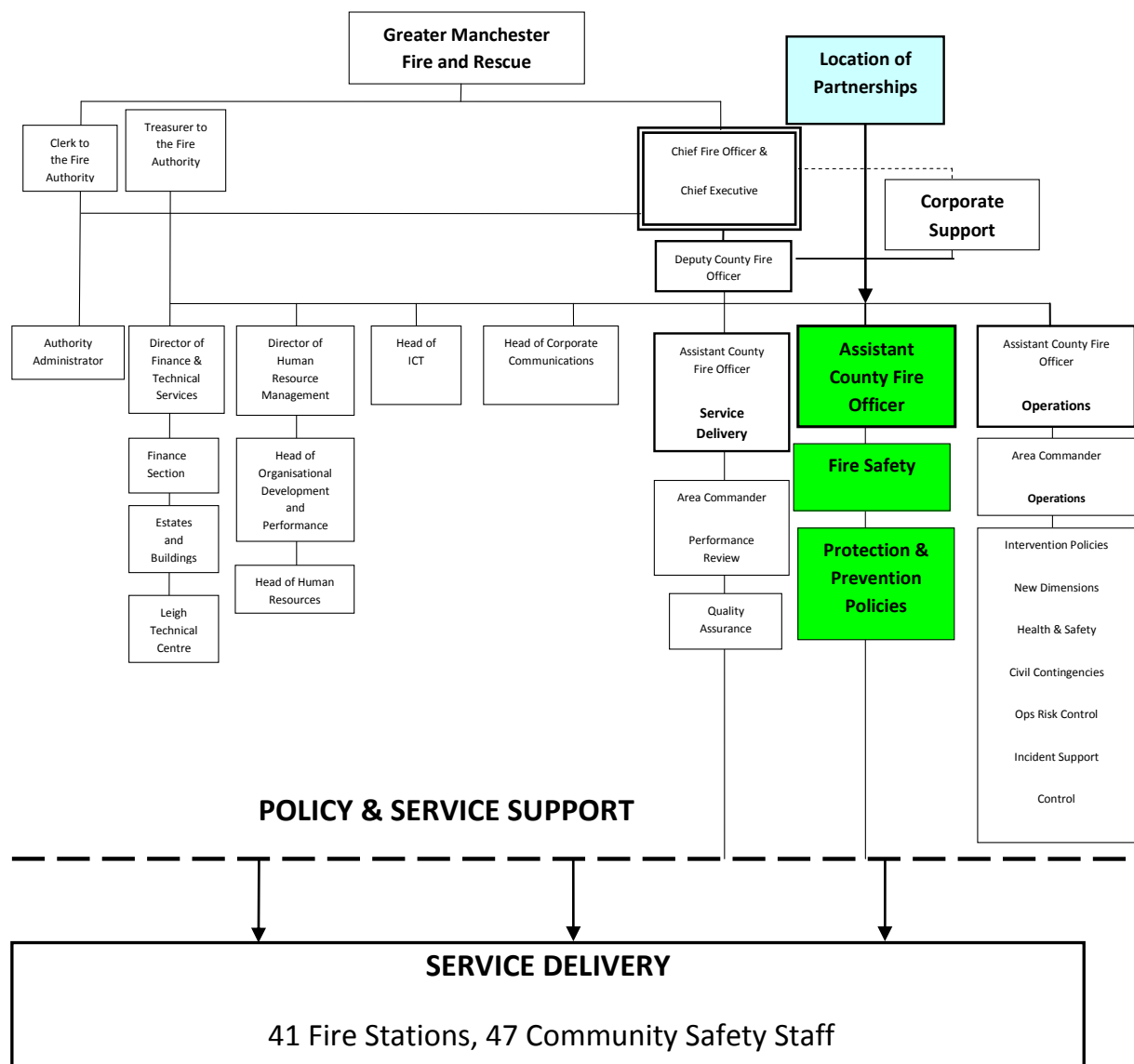
Figure 2: Greater Manchester Fire and Rescue Authority Structure

---

[1] This information was correct in 2008-2010 when the research was carried out. The organisation has since undergone changes to accommodate the changing focus of the organisation's activities.

3.3 Background to the Research Project

GMFRS had become aware that the impact and extent of influence of the newly developed community initiatives was not clear. There were also external demands to demonstrate these newly developed activities were good value for money, and delivering the intended results. Attempts were made to encourage evaluation, but no common agreement about evaluation practices was ever reached. The central services tried to impose an output evaluation model on the Boroughs, but were not successful in embedding evaluation into the community safety activities; hence the majority of the community interventions were never evaluated at sufficient depth. This resulted in GMFRS acknowledging that internal expertise in evaluation was limited and there was a need to develop a wider evaluation framework to guide and support GMFRS users in designing and conducting evaluations of the community safety initiatives (KTP Grant Application and Proposal Form 2007). GMFRS had very little previous experience in research or evaluation, hence a Knowledge Transfer Partnership (KTP) was formed with the University of Salford to address this gap; the University would provide the necessary expertise in identification, selection, and knowledge of how to apply relevant evaluative techniques, and how these can be tailored to the GMFRS organisational context. The aim of the KTP project was to develop an evaluation toolkit that was sensitive to the variability and complexity of GMFRS community safety initiatives and yet capable of being utilised by internal personnel with no specialist expertise in evaluation. This then developed into a PhD through which the researcher could further investigate evaluation capacity building in a large public sector emergency service.

GMFRS believed the development of evaluation tools would enable the organisation to identify and evaluate critical success factors within both existing and new preventative initiatives that would inform the design and delivery of future activities. As a direct result, the organisation would have better business continuity management helping to support more sustainable communities (KTP Grant Application and Proposal Form 2007). By evaluating interventions, GMFRS believed they would be better able to determine what works best, and therefore enable better direction of valuable resources (KTP Grant Application and Proposal Form 2007). The organisation had the following expectations of the research project:

GMFRS's aim for the project:

- To develop an evaluation toolkit, trial its application and effectiveness, and to embed evaluation and feedback into GMFRS public prevention initiative design process (KTP Grant Application and Proposal Form 2007:1).

GMFRS expected outcomes of the project (KTP Grant Application and Proposal Form 2007:8):

- Evaluation tools, techniques and protocols skills developed, including information on parameters of application; initiatives and interventions; and validation protocols. Embedding this within the organisation will enable GMFRS to have more effective planning and better business continuity management;

- Ability and enhanced understanding enabling GMFRS to further develop and update the systems to cope with complexity in initiatives and partnership initiatives, and identification of services to be offered to meet;

- Improved corporate and individual understanding of evaluation techniques, their importance within the planning process and the significance of evaluation to inform subsequent decision making;

- Improvement in the efficiency of internal resources and the development of staff to increase the visible expertise of the organisation;

- More systematic use of economic evaluation at the start of and during the life of each initiative, leading to informed decision-making and the most effective use of resources thereby building capacity; and improved economy and efficiency; and

- A further enhanced profile and reputation of Greater Manchester Fire & Rescue Service at a national level.

3.4 Locating the Context

The initial goals of this section were to examine what types of organisations seek to develop their internal evaluation capacity, and how they had experienced the ECB initiatives. The context is an important contemplation for future learning and development of the ECB field. Preskill and Boyle (2008) argue that ECB represents the next evolution

of the evaluation profession, and has the potential for transforming the field. But the field lacks empirical research which is required to build a more robust knowledge base (Sanders 2002, Cousins et al. 2004, Preskill and Boyle 2008). A survey of the current ECB studies confirmed the argument advanced by Sanders (2002), Cousins et al. (2004) and Preskill and Boyle (2008), that very limited amount of information exists about the organisational context of the ECB efforts. All of the ECB efforts were carried out by external evaluators, and concentrated on describing how the developer had carried out the ECB activities. The literature did not offer any explanations of how the organisations had reacted to or benefited from the ECB, and the extent to which the results of the efforts were sustainable. A dialogue between the organisations and the ECB developers would have captured both the organisational perspective and the development of the ECB model. However the lack of discussion about the links between the organisational context and the chosen ECB approach could be explained by the fact that much of the ECB literature presents theoretical descriptions of ideal ECB models which have not been tested empirically (for example Khan 1998, Stevenson et al. 2002, Barnette and Wallis 2003, Monroe et al. 2005, Forss et al. 2006, Huffman et al. 2008, Taylor-Powell and Boyd 2008). The literature review also identified that the models that had been empirically tested / developed (Minnett 1999, Porteous 1999, King 2002, Milstein et al. 2002, McDonald et al. 2003, Cohen 2006, Naccarella et al. 2007, Taut 2007, Adams and Dickinson 2010) had been written from the ECB developers perspective (apart from McDonald et al. 2003 which includes contributions from the organisational perspective) and as guides to other ECB developers/academics, thus containing little analysis of the relationship between the context, the type and extent of the ECB effort. This is an interesting finding because it was noted in the previous chapter that ECB definitions are context specific and constructed to suit the purpose of the work; however an examination of the contexts reveals that very

little is known about the organisations where ECB approaches are developed. In Table 3 the context of the few empirically tested models found in the literature is detailed.

Table 3: Summary of the organisational context of ECB efforts

ECB APPROACHES	
Author	**Context**
Minnett (1999)	Non-profit agency providing youth services. Values and encourages self-refection to enhance evaluation use and organisational learning. ECB effort received substantial support from funders.
Porteous (1999)	Public health organisation. Has an evaluation culture, but management did not have sufficient knowledge of evaluation, and evaluation has not been well integrated into the programme management cycle.
King (2002)	School district consisting of 40 schools. Internal programme evaluation is a relatively recent addition to district practice.
Milstein et al. 2002	The organisation engages in a vast array of evaluation activities and has several well-developed evaluation systems. But distribution of evaluation expertise is uneven, and evaluation is not practiced consistently across all areas.
McDonald et al. (2003)	Public sector organisation managing large and complex projects, competing for government funding. The organisation had a strong science research culture, with a large number of staff trained or experienced in approaches to research.
Cohen (2006)	A non-profit organisation promoting the use of technology in education. The small staff, primarily educators, was characterised by high energy levels and a can-do attitude. Programme managers chose to fund evaluation generously, allocating close to 15% of total grant funds.
Huffman et al. (2006)	K12 Schools. Schools need to develop evaluation capacity to manage and use the multitude of data they gather about students, to make decisions about how students can improve scores the next year.
Naccarella et al. (2007)	Government funded organisations providing health services to general practitioners.
Taut (2007)	International development agency that already had an evaluation department, but evaluation did not contribute to a learning culture at all levels of the organisation.
Compton (2009)	Nationwide, voluntary health organisation. Historically, the organisation has devoted little systematic effort to determining programme effectiveness. Few staff trained to do professional programme evaluation, but logic models used regularly, and organisation an evaluator.
Adams and Dickinson (2010)	Community and public health organisations that receive government funding. For more than 10 years, it has funded evaluators to train and support this workforce in developing evaluation and related skills.

A review of Table 3 reveals the following commonalities in the organisational contexts of the ECB approaches:

- The ECB efforts took place in either public sector or non-profit organisations which operate in the field of education or health;

- ECB models were developed by external evaluators;

- ECB efforts were long term, high cost interventions;

- The organisations were ready for evaluation, i.e. staff members had experience with research or evaluation activities, and the organisation was supportive of the ECB effort; and

- The organisations had either existing evaluation departments, external evaluation assistance, or experience with evaluations.

The key term that emerged from the literature was 'learning organisation'. According to Davidson (2001), this is a phrase used to describe an organisation that creates useful knowledge, disseminates it effectively, and uses this knowledge to improve organisational effectiveness. As a result of the learning culture they are better able to anticipate change, respond/adapt more quickly to change, and perform better and survive longer than organisations that do not learn so well (Davidson 2001). Those ECB efforts that provided some detail about the organisational context demonstrated that organisational readiness, a considerable amount of support for the ECB efforts, and a learning culture were the key factors to successful ECB. Minnett (1999), McDonald et al. (2003), and Naccarella et al. (2007) even argued that the organisational culture was the determining factor in developing successful ECB models. In addition Minnett (1999), Cohen (2006), Naccarella et al. (2007) and Adams and Dickinson (2010) mentioned successful ECB

efforts required a large financial investment and a substantial amount of resources. The only author describing a different ECB experience was Taut (2007) who argued that the organisational context of her ECB effort differed greatly from all the existing approaches. According to Taut (2007), the organisation had insufficient capacity for evaluation and a low-trust environment in which distrust caused resistance. However, the organisation already had an evaluation department in place, and for them ECB corresponded to improving the utilisation of evaluation findings for learning and development purposes.

GMFRS is a public sector emergency service that cannot be defined as a learning organisation, and even though they had some evaluation materials available for the staff to use, they did not have extensive experience of research or evaluations, either internally or externally conducted. GMFRS had no skills in data collection, and very limited understanding of the activities and processes associated with evaluation. GMFRS can relate to the external and internal ECB demands of the other organisations. GMFRS' need for ECB derived from internal improvements, and external demands to demonstrate effectiveness and improve accountability. However, the GMFRS ECB effort was not designed to affect the whole organisation, only a small part of it, as shown in Chapter 2. This particular area of the organisation dealing with the community interventions had very limited resources, money and time, for the ECB effort, hence the GMFRS experience differs from the other ECB efforts. This has encouraged the investigation of two questions in the thesis that had not been discussed in the literature:

1. How to develop ECB in an organisation that cannot be characterised as a learning organisation or that does not have an evaluation background?

2. Are there any barriers to developing ECB, and if yes, how to overcome those?

3.5 Aims and Objectives of the Research

As identified in Chapter 2, the literature has gone as far has detailing individual and organisational processes required for successful capacity building efforts, however insufficient attention has been paid to the usability of these models for non-learning organisations, and the embedding of the skills and processes into the daily activities of an organisation. The knowledge gaps identified in the literature review and the organisational needs have resulted in the formation of the following research aim and objectives:

Aim:

The ultimate aim of the study is to develop a theoretical model for the use of evaluation capacity building (ECB) in a large public sector emergency service.

Objectives:

1. to assess the process of:
   a) developing an evaluation framework and toolkit, and
   b) supporting the embedding of an evaluation culture in the GMFRS against the original aims and objectives of the project, and comparing the findings to a relevant theoretical framework;
2. to investigate the challenges and benefits of non-experienced personnel using the evaluation toolkit successfully; and
3. to assess the success of a project designed to develop a self-evaluation framework and toolkit for use by non-specialists in the GMFRS

3.6 Conclusion

In this chapter an introduction to the context of the research was provided. The chapter introduced the organisation, examined its role in the wider community, and discussed the organisational need for evaluation. The contexts in which the other ECB models have been developed were studied, and conclusions drawn about further gaps in the ECB literature. Finally, after a contemplation of the gaps in the literature and the organisational needs, aims and objectives were set for this action research study.

4. METHODOLOGY

This chapter contains a critical examination of the theoretical debates that dominate the world of social sciences, and provides a justification for selecting the methodology of action research, and the use of mixed research methods. The chapter is divided into six sub-sections. In the first one, the philosophical debates that form research paradigms -- the guiding principles of research -- are examined. In the second, the reasons for adopting a mixed method/pragmatic research paradigm are explained. The third section is a theoretical overview of the chosen methodology -- action research -- and in the fourth section the reasons for adopting this particular methodology are explored further. In the fifth, the researcher's role in the study is discussed, and the final section concludes the chapter by summarising its content.

4.1 Research Paradigms

Social researchers approach research problems from different theoretical and methodological perspectives. Characterising the nature of the link between theory and research is by no means a straightforward matter (Bryman 2008). The selection of an appropriate approach, the selection of a research paradigm, is based on the researcher's ontological and epistemological beliefs. Social science research should establish and understand the ontological and epistemological questions rather than take them as universal and obvious truths (Mason 2002). They are interlinked with the choice of methodology (as shown in Figure 3), which establishes how the researcher goes about finding what they want to know (Crotty 1998). However, it has to be noted that "the beliefs are basic in the sense that they must be accepted simply on faith; there is no way to

establish their ultimate truthfulness" (Guba & Lincoln 2003:21). Before proceeding to discuss the approach that was chosen for this research, all of the above mentioned concepts need to be further examined to establish the impact they have on research practices.
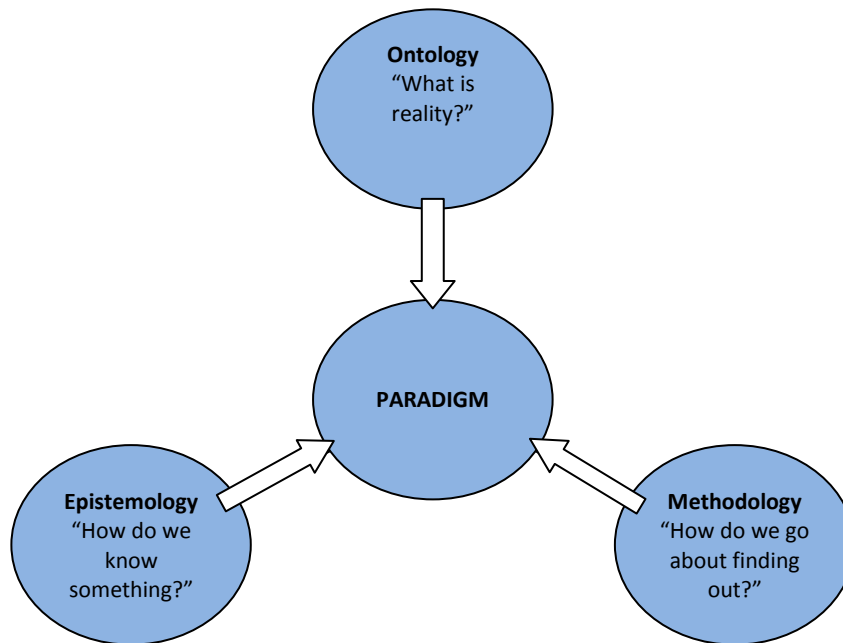
Figure 3: Founding Principles of Paradigms

Ontology questions the notion of social reality, the very nature and essence of things in the social world, by asking what is reality and what is there that can be known about it? (Mason 2002, Guba and Lincoln 2003, Blaikie 2007, Bryman2008). Ontological questions encourage the researcher to contemplate whether social realities are objective entities that exist independently without the thoughts and activities of human beings or whether they are constructed from their perceptions and actions (Guba 1990, Blaikie 2007, Bryman 2008). Epistemology, closely linked to ontology, questions what is regarded as

acceptable knowledge and asks what does one count as knowledge of social things (Blaikie 2007, Bryman 2008). Epistemology guides the researcher to question how human beings come to have knowledge of the reality that they believe to exist. It provides the grounds for establishing if knowledge is produced by the interaction of the human beings that construct reality, which is then interpreted by the researcher studying those particular constructions, or if it arises from a set of natural laws that reality provides, and which the researcher observers to describe the phenomena that we experience (Trochim 2006).

Paradigms are research strategies that are located within the broader frameworks of theoretical or philosophical perspectives i.e. the ontological and epistemological considerations. Paradigms establish the limitations of the inquiry, guide the way the research questions are set, and constrain the way the questions are answered (Blaikie 2007). It was previously argued that linking these philosophical questions to research practice is a complicated matter (Bryman 2008). Guba (1990:17) defines a paradigm as the "basic set of beliefs that guide the action", but argues most people asked to define the word paradigm are unable to offer a clear statement of its meaning. "Thomas Kuhn who brought the term in to our collective awareness has himself used the term in 21 different ways" (Guba 1990:17). In the research method literature, paradigms are also knows as philosophical assumptions, epistemologies and ontology (Crotty 1998, Betzner 2008) schools of thought (May 2001) and worldviews (Creswell 2003). May (2001:8) questions if the lack of consistency undermines the idea of scientific discipline; but argues that the assumption that "science is an all-embracing explanation of social world" has to be challenged. The beliefs researchers bring to research have continually evolved over time (Creswell 2003), and as the definitions of the term 'paradigm' are so fluid they allow researchers to reshape the term as understanding of it improves (Guba 1990).

The prevailing dispute among the dominant paradigms: positivism and interpretivism, is often known as the "paradigm war" (Guba 1990). These principal paradigms represent the opposing sides of the research paradigm spectrum and vary in their epistemological, ontological, and methodological approaches, as shown in Figure 4.

Positivism
Ontology: Realism
Epistemology: Empiricism
Methodology: Quantitative
Orientation: Deductive

Interpretivism
Ontology: Idealism
Epistemology: Constructionism
Methodology: Qualitative
Orientation: Inductive

Figure 4: Opposing paradigms

Positivism represents the traditional science based approach to social research. Positivists are ontological realists: reality is external to the activities of human being and understood in terms of natural laws (May 2001). Human beings are products of the environment and the causes of human behaviour are regarded as being external to the individual (Bryman 2008). The key to the conduct of positivistic research is objectivity. The researcher employs an external position, and controls nature and methods that place the point of decision with nature rather than the inquirer (Guba1990). Learning is derived by analysing experimental experiences, and concepts and generalisations are summaries of particular observations (Blaikie 2007). Positivism is traditionally aligned with quantitative methods and data types (Betzner 2008), and deductive theory where hypotheses are drawn from theory and then subjected to empirical scrutiny (Bryman 2008).

Interpretivism -- a general term for qualitative research that aims to interpret human experience (Blaikie 2007) -- also known as constructivism (Creswell 2003), provides a more human centred approach to social science and aims to challenge the core beliefs of positivism. Denzin and Lincoln (2000) argue positivistically based quantitative researchers believe the keys to the conduct of real social science to be objectivity, distance and control, however controlled situation are not always representative of social reality because they require the manipulation of social spaces. Interpretivism challenges the single reality worldview and argues that reality is socially constructed and exists only in our thoughts; hence there is not only one but multiple realities (Guba 1990, Blaikie 2007, Creswell 2003). Human beings interpret their realities and give meaning to them. Knowledge of things is gathered by combining the different constructions human beings have of the same topic. In contrast to the positivistic objectivism, interpretivist research is subjective. Reality has to be discovered from the inside rather than being filtered through an outside expert's theory (Blaikie 2007). The researcher's views, past experience, and constructions are also part of the research process and the study, the stories voiced represent an interpretation and presentation of the author as much as the study (Creswell 2003). Interpretivist approaches are often associated with qualitative methods and data types (Betzner 2008), and carried out in an inductive manner where the researcher starts with empirical questioning and then focuses on feeding the findings into a theory (Bryman 2008).

Within this section of the chapter the philosophical questions that dominate the world of social science have been examined. Knowledge of the formation of the paradigms is essential, as ontological and epistemological questions offer useful considerations and a practical and inquisitive framework for the researcher of social matters. However the

positivistic/interpretivist dichotomy creates a restricting framework that shapes the way research is carried out (Johnson and Onwuegbuzie 2004, Plano Clark and Creswell 2008). The paradigms limit the way a researcher can examine the research questions, and if only one approach has to be employed, there is a risk that some aspects of the subject can be left undiscovered (Mason 2006). If the research paradigms shape the choice of methods then the researcher is likely to rule out particular methods from the start, and not be governed by the research process and the context as it unfolds (Brannen 1995). The subsequent chapter will examine a mixed method/pragmatic research paradigm which focuses on the research outcomes rather than the philosophical questions as the driving force of research.

4.2 Mixed Method/Pragmatic Research - "the Third Paradigm"

Leading on from the previous discussion, inquiry paradigms define what the inquiries are about, and what falls within and outside the limits of a legitimate inquiry (Guba and Lincoln 2003). The positivist/constructivist dichotomy encourages people to situate their social science activity on one side of it or the other, which hinders the development of meaningful social theory and explanation (Mason 2006:15), because it alters the focus of the research process from the actual outcomes to a philosophical debate. The "paradigm war" has focused on the differences between positivist and constructivist orientations, hence placed limitations on research practices, and the selection of research methods. The debate has guided researchers to certain methods and ways of linking theory to research, which appears very restrictive, as argued by Mason (2006:10): "social scientists fail to see the worlds of experience and understanding if they define research as purely qualitative and/or quantitative, as though those categories and that division encapsulate all we are capable of knowing". Rather than dismissing the others' work based

on wholly contrasting assumptions, Morgan (2007) suggests that our goal should be to search for useful points of connection in the paradigms. Not challenge or dismiss the existing paradigms but rather aim to embrace their strengths and use them in an overlapping way (Creswell 2003). Hence authors such as Patton (1997), Tashakkori and Teddlie (2003), Johnson and Onwuegbuzie (2004), Bryman (2006), Mason (2006), Morgan (2007) and Plano Clark and Creswell (2008) have come to promote the use of a "third paradigm", which challenges the restrictive dichotomy by promoting the use of whatever methodology suits the situation, and draws on the best principles of qualitative and quantitative inquiry. This "third paradigm" is known as pragmatism or mixed method research (Plano Clark and Creswell 2008). Mixed method research is founded on the philosophy of pragmatism (Johnson and Onwuegbuzie 2004, Plano Clark and Creswell 2008); however as authors such as Patton (1997), Johnson and Onwuegbuzie (2004) and Mason (2006), associate positivistic research with the use of purely quantitative methods, and interpretivist research with the use of purely qualitative methods, pragmatism has also become synonymous to mixed method research. Pragmatism is a paradigm, which allows researchers to mix qualitative and quantitative methods, and the worldviews that underpin them.

The pragmatist research paradigm is not committed to any one philosophy or reality because the main concern lies in outcomes of research; hence the researcher has the freedom to choose the methods, techniques, and procedures that best meet the needs and purposes (Creswell 2003, Johnson and Onwuegbuzie 2004). This gives a new kind of freedom to researchers, and provides them with a "capacity to explain, and to ask and answer rigorous and useful questions in our complex social environment, we need to understand how contexts relate to social life, and factor this understanding into our

explanations" (Mason 2006:15). For a pragmatist, reality is the moment, and truth is what works at the time. Ontology has room for mental and social reality as well as the more micro and more clearly material reality (Creswell 2003), as demonstrated by Johnson and Onwuegbuzie (2004:16) : "It is not a matter of opinion (or individual reality) that one should or can drive on the left-hand side of the road in Great Britain - if one chooses to drive on the right side, he or she will likely have a head-on collision, at some point, and end up in the hospital intensive care unit, or worse. This is a case where subjective and objective realities directly meet and clash". Betzner (2008) also notes the challenges of using purely positivist or interpretive approaches in real life settings. She argues that when investigating complex phenomenon, such as community initiatives or policies, the use of interpretive approaches might be more effective. However, in these settings, interpretive methodologies also face significant challenges, as the impact of large scale interventions is often too extensive to make the sole use of qualitative approaches (Betzner 2008).

When it comes to orientation to research and data collection, the pragmatic paradigm relies on a version of abductive reasoning that moves back and forth between deduction, associated with positivism, and induction, linked to constructivism, by converting observations into theories and then assessing these theories through action (Morgan 2007). Data is treated as unique to time and place and there is no specific question or hypothesis to be tested in terms of causation or correlation, in some situations the qualitative approach will be more appropriate, and in others the quantitative approach more suited (Johnson and Onwuegbuzie 2004). Each way has strengths and weaknesses; qualitative data offers detailed, rich description, capturing variations between cases, and quantitative data facilitate generalisable comparisons (Patton 1997). Qualitative researchers are motivated to understand the world through their own eyes (Hesse-Biber

and Leavy 2003), in an environment that is natural to the research subject. Qualitative methods often bring to light profound insights from personal experiences, which cannot so easily be detected from quantitative and standardised methods of inquiry (Dick 1993). However, the strengths of positivist/quantitative research are precision, generalisability, reliability, and replicability (Betzner 2008). They also share similarities. Both quantitative and qualitative researchers use empirical observations to address research questions, and both sets of researchers incorporate safeguards into their inquiries in order to minimise confirmation bias and other sources of invalidity (Johnson and Onwuegbuzie 2004). They also share a consensus about which questions are worth asking and which methods are most appropriate for answering them (Morgan 2007).

Within this section a justification for the use of pragmatic/mixed method paradigm to research has been presented, and an argument advanced that, rather than dismissing and/or challenging the existing paradigms there ought to be an acknowledgment of their existence and importance in establishing and understanding the concepts. Both quantitative and qualitative research approaches are important and useful. The goal of mixed methods research is not to replace either of these approaches but rather to draw from the strengths and minimise the weaknesses. For this research, the pragmatist paradigm offered a useful middle position, philosophically and methodologically, and did not place limitations to the questions asked. The next section contains a discussion of the methodology chosen for this research -- action research -- which also focuses on the outcomes of research and makes use of multiple research methods to suit complex social situations.

4.3 Action Research as a Methodology

Action research does not have a coherent history, but many authors trace its origins back to the social experiments of Kurt Lewin in the 1940s (Reason and Bradbury 2001:2), and have adopted some version of Lewin's spiral steps composed of circles of planning, action and fact-finding about the result of the action (Winter and Munn-Giddings 2002:10). Rather than aiming to give an explicit definition of action research, authors seem to be more comfortable with describing what it involves and how it generates knowledge. All action research approaches involve the process of change, but ask different questions, depending on the author's perspective (McNiff and Whitehead 2006). Action research has both practical and theoretical implications; hence the descriptions often include a notion of action and change (Baskerville 1999, Reason and Bradbury 2001). Baskerville (1999) describes action research as a widely used and well established research method which produces highly relevant research results due to the practical action aspect aimed at solving a problem situation while carefully informing theory. The dual commitment of the methodology contributes both to the practical concerns of people, by changing some aspects of their social environment, and to the theoretical understanding by developing the existing understanding of the subject (Baskerville 1999). McNiff and Whitehead (2002:13) see it as a process of learning from experience, "a dialectical interplay between practice, reflection and learning", whereas Somekh (2006) argues that in addition to analysing, describing and theorising social practices, it involves working in partnership with participants to reconstruct and transform practices. Baskerville (1999) claims the social world is best studied by introducing changes into social systems and observing the effects of these changes. The use of other methodologies should not be dismissed in studying social change, but action research has an important role in applied research. Action

research is an effective way of gaining knowledge in situations where the researcher has the ability to act as the driving force for change, observe the effects, and develop further knowledge about the subject under investigation. The aim of action research is to understand and transform, rather than explain, predict and control, which is essential in understanding a social system and discovering opportunities for change (Dick 1993). McNiff (2000) also argues that action research is effective in improving participants' lives, as it can improve the quality of their own learning, which has the potential to influence the lives of others for good.

Dick (1993) argues action research should be seen as a collage of attitudes, frames of mind and orientations to inquiry, rather than as a precisely defined set of methods. The methodology is representative of the unknown situation under investigation; hence, the research process and the research questions can appear vague at the beginning. Action research cannot be started with specific research questions as the flow of the inquiry will determine them (Dick 1993), and the solutions to the research activities have be to found from the context (Stringer 2007). Even though action research methods and inquiry style are constructed to a particular situation, it has to be noted that action research follows a methodological structure that relies on testing, and on a theoretical framework to support the claim of knowledge. In essence action research is "learning by doing" and, as Figure 5 shows, involves the stages of identifying a problem, taking an action to resolve it, evaluating the usefulness of the efforts, and if not satisfied, trying again (O'Brien 1998). In the cyclical process, data are gathered and peer reviewed to demonstrate the reality of the practice and its potential impact on others. Knowledge is gained through studying the researcher's and the research participant's experiences, and learning used to inform new practices. Research outcomes are seen as new starts rather than closures (McNiff, 2000).
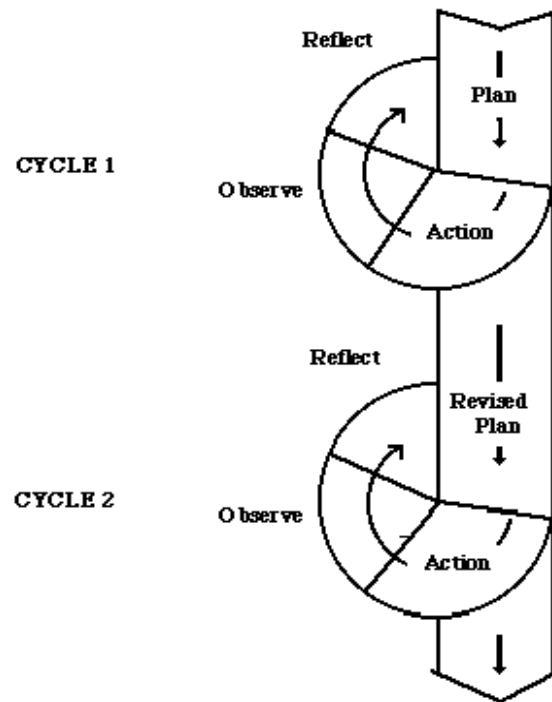
Figure 5: Action Research Model (MacIsaac 1995)

Action research draws on many ways of knowing and there can never be a right way of doing action research. To ensure action research is a rigorous method of inquiry and meets necessary quality standards, the researcher needs to be transparent about the research process and articulate all the actions and findings to the participants and the wider world (Reason and Bradbury 2001). Validity, reliability, and creditability are measured by the willingness of local stakeholders to act on the results of the action research (Denzin and Lincoln 2000, Levin and Greenwood 2001) and the degree by which the conclusions are supported by the analysis (Mason 2002). Action research needs to demonstrate it offers explanations rather than just observations and descriptions of practice (McNiff and Whitehead 2002).

Within this section a theoretical overview of the methodology of action research has been given. The key issues raised are that action research a) is driven by change, b)

relies on both deductive and inductive ways of linking research to theory, c) sees reality as the current state of affairs, and d) draws on many ways of knowing. All of these issues will be examined and discussed in further detail in the next section that explains the reasons for adopting this particular approach, and why this approach was chosen over other methodologies.

4.4 Why Action Research?

4.4.1 Focus of Change

The organisational situation was highly complex at the start of the research. As discussed in Chapter 2, GMFRS was going through a modernisation programme, and had adopted a new business model to improve performance. The new model presented two new demands to the organisation: the development and delivery of numerous public protection initiatives, and the need to demonstrate that these newly developed activities were good value for money and delivering the intended results. There was a fair amount of resistance to the delivery of the interventions because of the additional workload they created, and, due possibly to GMFRS not being a learning organisation, their attitude to evaluation and assessment was very negative. Even though the research project was not very popular at the start, the organisation still needed a new evaluation framework to help provide evidence of the impact of the initiatives. Action research was chosen because its responsiveness to challenging situations, and because of its emphasis on action, change and outcomes of research. Action research allows the main emphasis to be on action, with research as a fringe benefit (Dick 1993); action research simultaneously assists in practical problem solving and expands scientific knowledge (Baskerville 1999). The arguments that

action research cannot be started with specific research questions as the flow of the inquiry will determine them (Dick 1993), and the solutions to the research activities have be to found from the context (Stringer 2007), suited the context as the organisation was struggling to demonstrate the impact of their activities but did not know why. The research questions and processes were not clear from the start, and the cyclical and reiterative process of action research allowed the researcher to investigate assumptions, and try again.

4.4.2 Democratic Methodology

The project aimed to develop a process suitable to the organisation and its staff members' needs; hence a democratic methodology, that involved the staff in the research process, was required. The argument advanced by Denzin & Lincoln (2000), that local people have the specific knowledge to create social change, whereas the researcher brings the theoretical and methodological knowledge, suited the situation perfectly as the product of the research, an evaluation toolkit, was designed to be used by the GMFRS personnel. It was essential to include the staff in the research process to ensure the product was suitable purely for their skills. Involving staff in the research also acted as a way of disseminating findings, and ensured the skills and knowledge gained during the research project were embedded in the organisation. The primary focus of action research is on turning the people involved into researchers, and helping them to apply the learning to real life situations (O'Brien 1998). Therefore the methodology also supported all the roles that the researcher had taken on -- a researcher, a research participant, and an employee. The methodology allowed the researcher to enhance others' learning, as well as record personal development, as in action research the researcher becomes part of the study (Baskerville 1999, McNiff and Whitehead 2006).

The organisation did not have a readymade user group for the evaluation system that could have provided constant assistance with the research, and that the research project could have empowered to take evaluation even further in the organisation. Everyone involved in evaluation at the time of the research was engaged in the research process, either testing the toolkit or providing feedback. A communications plan (Appendix A) was also created to keep all members of the organisation informed. The plan utilised a good variety of communications methods from a blog to attending meetings, to allow everyone an opportunity to get involved with the research. In the thesis several references are made to a "project core team". The core team consisted of an academic from the University of Salford, and three GMFRS staff members, two uniformed and one non-uniformed. The academic provided guidance on research related issues, whereas the GMFRS officers acted as mediators between the researcher and the organisation, organising contacts and introducing the researcher to the organisation. Even though the researcher was a full time employee for the duration of the project, due to organisational changes did not automatically fit under a specific organisational structure or department. The officers were in roles where they could influence people across the organisation, and had the power to influence the delivery and the outcomes of the research project. The core team made all the key decisions regarding the research project, and assisted in planning the direction of the ECB activities.

4.4.3 Flexibility

The final aspect that supported the choice of action research as the most suitable methodology was its flexibility. As the attention is on outcomes and change, it was not linked to certain methods and/or paradigms, but rather on what the situation needed.

Action research accommodated and complemented the use of mixed method research, as a paradigm and as orientation to research. It also works by inducting and deducting theory, and allows the flexibility to employ both a qualitative/interpretivist approach, as the researcher makes the value choice of pursuing situation specific knowledge rather than generalisable knowledge, and also a positivistic/quantitative structure, by developing theories that inform the actions (Swepson 1995). The paradigms underpinning other research methodologies created an either-or dichotomy, which was too restricting for applied research as they were either lead by a hypothesis or a philosophical position. A purely positivistic/inductive/quantitative research approach, such experiments and quasi-experiments, would not have been suitable for the study because the organisational processes have to be taken as the reality, and the people's views of the processes have to be treated as constructions of the reality that they live in. The situation was not controllable to the extent required for a positivistic/inductive/quantitative research approach and relied heavily on subjective data. For the success of the project, it was essential to include staff views and experiences in the study. Purely interpretivist/deductive/qualitative approaches, such as grounded theory, ethnography, or phenomenology, did not provide the action orientated framework that would have supported the pragmatist position to research. In qualitative forms research is the primary focus and action is often a by-product - more attention is given to the design of the research rather than to other aspects (Dick 1993). Purely participatory methodologies, such as participatory action research, were also felt unsuited because the research project did not have the necessary resources; the organisation did not have a readymade user group for the evaluation project that could have benefited from this kind of approach − as described in the previous section.

Within this section the rationale for choosing the methodology of action research was provided. In essence its practical application, flexible nature, and focus on solving practical problems were the key characteristics that guided the decision. Action research focused on the implementation of change, supported the involvement of GMFRS personnel in the research process, and complimented the use of mixed research methods. The next section explains how the methodology was used during the research process, and provides a diagrammatic presentation of the action research Cycles of this study (Figure 6).

4.5 My Action Research Cycles

The study was divided into three Cycles (Figure 6). The overall aim of the various activities carried out during the Cycles was to develop and embed an evaluation framework for GMFRS. Each Cycle consisted of a planning stage, in which actions were chosen for the Cycle, based on prevailing and pressing issues. After acting upon research problems, an observation stage followed where the findings of the action stage were assessed against relevant literature. The final activity of each Cycle was to reflect on the success of the Cycle, the research activities, learning and usefulness of literature. This section of the chapter provides an overview of the content and a diagrammatic presentation of the Cycles; the Cycle chapters will provide a more detailed account of the research activities carried out.

4.5.1 Cycle 1

The aim of the first Cycle, which ran from April 2008 to December 2008, was to understand the organisation's evaluation needs. The research activities included an examination of the community initiatives portfolio, a study of the initiative management documentation and the existing evaluation framework, and meetings with GMFRS personnel.

4.5.2 Cycle 2

The second Cycle, with the aim of developing new evaluation materials for GMFRS, was carried out between January 2009 and December 2009. The Cycle utilised a document analysis, a focus group and interviews to aid the format development of the new evaluation toolkit, and the examination of barriers to internal evaluation practices at GMFRS.

4.5.3 Cycle 3

The aims of the third Cycle, which was conducted between January 2010 and May 2010, was to develop a wider evaluation framework and embed evaluation into the organisation. The research activities included an examination of the process of developing recommendations for a wider evaluation framework, a focus group, and interviews with GMFRS staff to review the final version of the toolkit. The first two Cycles are very research focused, whereas the third Cycle concentrates on the change aspect of action research, and provides evidence of the transformation brought about by this action research project.
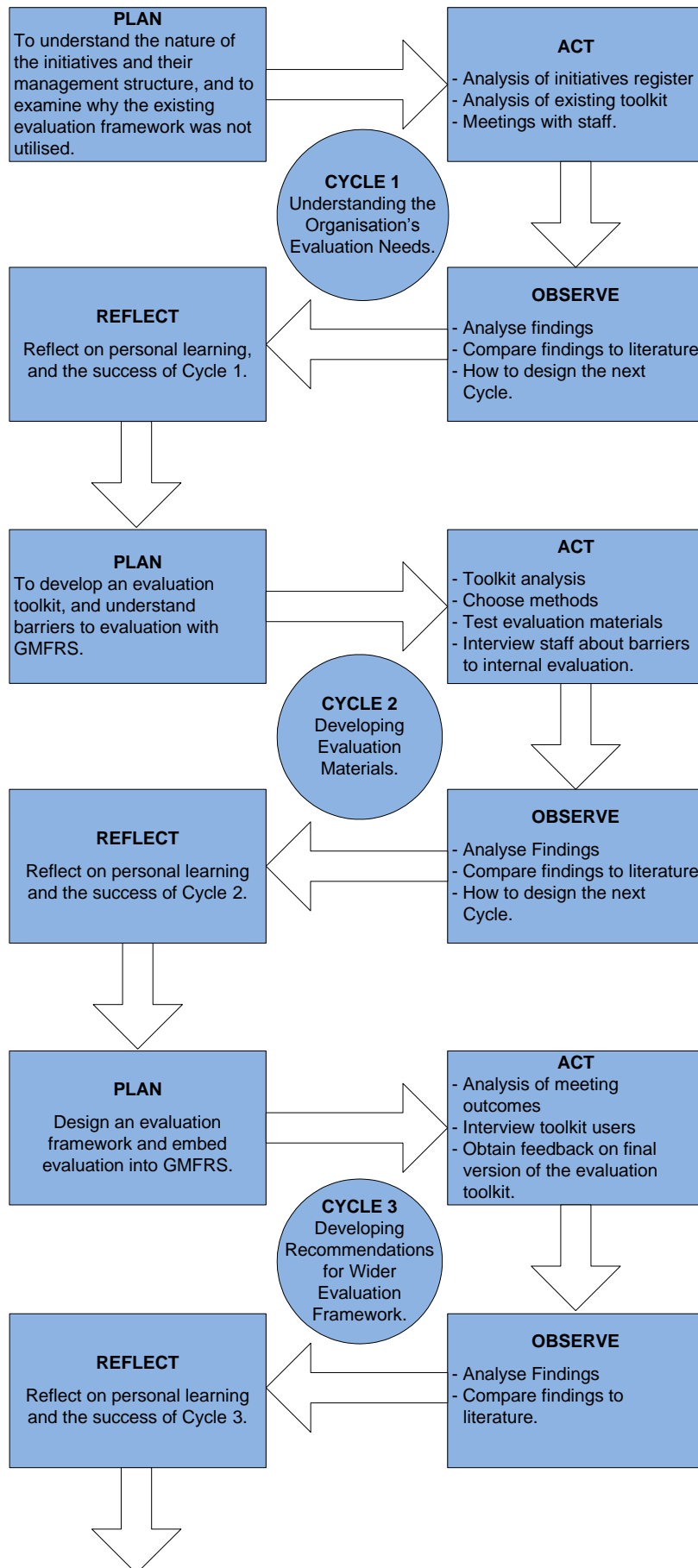
Figure 6: Action Research Cycles

4.6 Researcher's Background and the Role of Researcher

The researcher was fairly new to evaluation at the start of the research project. An interest in research and research methods originated from previous studies, which were further deepened by research and evaluation activities carried out for a charitable organisation. It could not be said that the researcher was passionate about evaluation, as it just seemed a way of collecting information about an activity. A familiarity had been gained about the concept of evaluation, and data collection methods, but not about the variety and scope of all different types of evaluation methodologies, or how they could transform projects and activities. This research project has completely transformed the researcher's view of the subject, and helped to understand that evaluation is much more than just an activity of data collection. It is now seen as a project planning and management tool, and an essential part of designing and running successful projects.

The research took place in real life settings, in an organisation where the researcher had been employed to manage a multi-agency partnership to develop an evaluation toolkit. The environment was not just a research site but also a workplace, henceforth in addition to the research activities, many other activities were carried out simultaneously. For the duration of the research a very complex role of a researcher, a research participant, and an employee was adopted to drive the change process. The roles did not vary much in their activities, but mainly in their relationships with the other staff members i.e. other research participants. The relationships were mainly collaborative – the researcher was building something solely for them and with them. The role of a research participant was also acquired, because the evaluation framework was developed for staff with no previous experience of systematic evaluations. As a PhD researcher the researcher sometimes felt

slightly distanced from the organisation and other staff members, because data was collected for personal purposes, to advance personal learning, and to be able to report the findings of the study to wider audiences in the form of this thesis. However keeping a very open relationship with the organisation via good communications channels was a key in maintaining good relationships with other staff members.

4.7 Conclusion

In this chapter the theoretical debates that dominate the world of social sciences have been examined, and a justification provided for selecting the methodology of action research, and the use of mixed research methods. The key arguments made in the chapter are: a) traditional research paradigms limit the way a researcher can examine the research questions, b) the mixed method/pragmatic approach to research focuses on the research outcomes rather than the philosophical consideration as the driving force of research, and provides a more flexible framework for applied research, c) action research is a context bound, action driven methodology that addresses real life problems, where participants and researcher collaboratively generate knowledge (Levin and Greenwood 2001). Within the chapter, it has also been demonstrated how the chosen paradigm and methodology (and selection of methods) complement each other. A description of the study Cycles was also provided, with a diagrammatic representation of the study, and a discussion about the role of the researcher. The next chapter of the thesis will review the data collection methods used in the research.

5. METHODS

In the previous chapter, a justification was provided for the chosen research paradigm, pragmatism, and the methodology - action research. Pragmatism is not committed to any one philosophy or reality as the main concern lies in the outcomes of research (Creswell 2003), hence it supports the use of action research and mixed method research. In essence, mixed method research refers to studies that utilise more than one method or paradigm -- a researcher can incorporate a variety of qualitative and/or quantitative research methods or analyses in the study, or just transform data through another approach (Tashakkori and Teddlie 2003). In Chapter 3, the challenges of utilising purely quantitative or qualitative research methods in real life settings and in applied research were discussed, and an argument advanced that complex social environments need more flexibility in the selection of research approaches and methods. Action research is based on learning by doing, and does not start with specific research questions because the flow of the inquiry will determine them (Dick 1993). The solutions to the research activities are found from the context (Stringer 2007), and the researcher has to be able to respond to the unknown situation by keeping an open mind regarding the use of research methods. In this section, an introduction is provided to the chosen methods, and a more detailed description of each method and their analysis is provided in the Cycles. The chapter is divided into eight sections: in the first four, the methods used in this study are discussed, in the fifth, a justification for the reflective model is provided, and in the sixth section, the note taking technique is explained. The penultimate section discusses the data analysis technique, and the final one concludes this part of the thesis by summarising its content.

A number of research methods, as shown in Figure 7, were used throughout the Cycles. Cycle 1 utilised document analysis, in Cycle 2 the researcher employed document analysis, observations, and interviews; and during Cycle 3 document analysis, interviews, and focus group were used. The terms quantitative and qualitative can refer to methods or data; in this thesis research methods are defined as techniques of data collection, and data is what is produced by particular types of methods (Blaikie 2007). Using Bryman's (2008) categorisation of research methods, all the methods, apart from document analysis, are categorised as qualitative. Qualitative research methods capture the research participants' or the researcher's views in their own words through interviews and observations (Patton 1997, Blaikie 2007); and quantitative methods are concerned with counting and measuring aspects of social life, hence qualitative data is often presented in words and quantitative with numbers (Blaikie 2007). In mixed method research, the researcher incorporates multiple methods in the study and transforms data from one approach to another (Tashakkori and Teddlie 2003). In this study, some of the qualitative interview data have been transformed into numbers/quantitative data, to make it easier to utilise, and to satisfy the information needs of the organisation. The organisation relied mainly on quantitative data and found it easier to use and process.
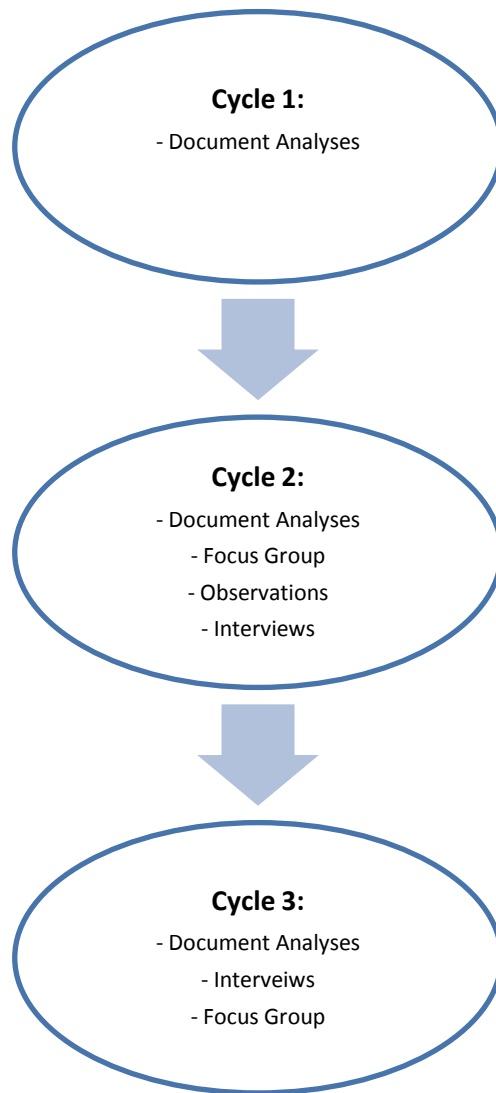
Figure 7: Research Methods

5.1 Document Analysis

Document analyses of various GMFRS policy and procedural guidelines that steer the organisation's youth intervention and evaluation activities were carried out throughout the study. The documents analysed related to the content and delivery of the initiatives: GMFRS initiative records, evaluation reports, meeting notes, and minutes of meetings. The document analysis was an opportunity to investigate nonverbal descriptions of the

organisation's culture, the initiative's delivery methods, and the way the organisation responded to recommendations made during the project. The documents that guide the organisation's activities and act as a driving force for their community interaction efforts also gave an insight into the factors influencing GMFRS's evaluation needs.

It is important to note that, in this thesis, the term document refers to all written materials studied during the research which was conducted between April 2008 and May 2010. Lincoln and Guba (1985:277 in Denzin and Lincoln 1994) argue that it is important to distinguish documents from records on the basis of whether they were prepared to attest some formal transaction. They claim records are produced for official use, whereas documents are for more personal reasons. Scott (1990:6 in Bryman 2008) also makes a distinction between official and personal documents by using the document's authenticity, creditability, representativeness, and meaning as the defining criteria. None of the documents analysed in this research could be defined as personal records, hence the defining criterion was not used.

Documents can provide very high quality information (Hodder 1994, Sarantakos 2005); however they can also hide a multitude of issues and guide a researcher away from any areas where, for whatever reason, they are unwanted. It is tempting to assume that documents can reveal something about the underlying social reality, and that the documents an organisation has generated are fully representative of what goes on in there (Bryman 2008). However, it is important to remember documents should always be examined and understood in terms of the very specific context in which they were produced (Atkinson and Coffey 2004 in Bryman 2008, Hodder 1994). Also, accessing documentation can be very challenging in a large organisation, though the effective use of

document repositories could be considered as a topic for another piece of research and will not be discussed in this thesis. From the outset, it was acknowledged that all the GMFRS documents were produced for a particular purpose, and consequently there were no guarantees of the quality of the data, but they provided interesting background information about the community initiatives. The contents of many of the documents, especially the initiatives register, used during this research raised more questions than provided answers, and that prompted the researcher to investigate the issues further. The documents gave a good indication of the organisational need for evaluation, and provided some context for the study.

5.2 Interviews

Interviews were used in Cycles 2 and 3. This method of inquiry was chosen because it provided opportunities to engage with participants and stakeholders on a more personal level, as well as understand their constructions of their roles in the organisation's evaluation process, and to examine attitudes to existing evaluation practices. The task of a qualitative interviewer is to provide the participants with a framework that allows them to express their views and experiences of events (Sewell 2009). The use of semi-structured questions as a starting point for both sets of interviews created a flexible structure that provided an opportunity to also explore unexpected issues, and the participant's constructions of their roles in the evaluation process. Interviews allow the researcher to engage the participants in an interactional exchange of dialogue, in a relatively informal setting (Mason 2002). Therefore, in addition to the flexible question structure, the researcher aimed to create a comfortable atmosphere, that the interviews could almost be seen as (one sided) conversations where the interviewees could freely discuss their views.

As the method eliminates peer influence, often associated with group methods, it was noted that the interviewees were very open about their views and willing to talk about unexpected and, maybe, even controversial issues. Though, as the researcher was fairly new to the organisation, it might have also influenced their perception of the situation: one of the participants noted outside of the forum of the formal interview that they would not have felt comfortable talking about their views to another member of staff.

Fontana and Frey (1994) note a growing number of scholars believe that most traditional in depth interviewing is only a way of manipulating people. Interviews can introduce a potential for interviewer bias, as the researcher can influence the situation by expressing (verbally or non-verbally) views of the topic and/or the interviewee, which can have a big impact on the participants answers. Social desirability can come to question if the interviewees feel they need to agree with the interviewer or answer in a "socially acceptable" manner (Bryman 2008). The interview situations were informal, however, they were never treated as an opportunity for an exchange of ideas. A conscious effort was made not to express any views of the topics, or comment on the participants' views. The researcher's role was to listen to their views and act as an enquirer. The aim of both sets of interviews was to understand the participants' constructions of the evaluation practices of the organisation, in order to improve them by developing a product and processes that would be suitable to their needs.

5.3 Observations

Both participant and non-participant observations were carried out during the research, the former in Cycle 2, and the latter in Cycle 3. In participant observations, the

observer joins a group and observes their activities, whereas in non-participants observation the researcher simply observes the activities and does not take part in them (Abbott 2009). The non-participant observer studies the subject from inside but employs a role that is clearly defined and different from the research subjects (Sarantakos 2005:220). The first two observation events were participatory in nature and took place during the second Cycle. The first one was an observation of the group dynamics of a focus group organised to choose research methods for the new evaluation toolkit, and the second observation was carried out at the toolkit testing stage to investigate how the staff members were conducting evaluations and how they collected data whilst still running the initiatives. The third observation was carried out in Cycle 3, when the researcher observed a focus group discussing the final layout of the evaluation toolkit. This was a non-participant observation, with the aim of collecting feedback about the evaluation materials.

Observations can give the researcher an alternative viewpoint of the research subject. They can reveal issues that would be difficult to discover otherwise, as well as complement information obtained by other techniques. Robson (2002) claims that interview and questionnaire responses are notorious for discrepancies between what people say and what they actually do. Even though observations allow the researcher to see what the participants actually do, the interpretation is always open to observer bias and vulnerable to selective perceptions and memory (Sarantakos 2005). The observer has very high control over the issues he/she wants to see, as they can place themselves in a particular location and choose to concentrate on a certain group (Adler and Adler 1994). The major advantage, as well as disadvantage, of observations is their directness. The researcher can watch what people do and listen to what they say, but by doing this they can influence their behaviour and communications in a way that has major impact on the

findings of the observations (Robson 2002). Based on this guidance, a very visible observer role was employed in which the purpose of the activity was made very clear. The reason for this was the audience. As a member of staff, my role involved developing a product with them and for them, so to maintain trusting relationship it was essential to remain open about my intentions. Any secrecy would have jeopardised my relationship with the participants, and had a major impact on the entire project. The potential for my observations to cause distractions to the initiatives was somewhat limited as they have well-structured delivery methods and guidelines the trainers have to follow to ensure consistent delivery. They would have had to carry out all the same tasks whether they were being observed or not. The observation findings have also been complemented with participant interviews.

5.4 Focus Group

Focus groups involve organised discussions with a selected group of individuals to gain information about their views and experiences of a topic (Gibbs 1997). For the purpose of this research, focus groups were used both as a research method as well as a decision making tool, in Cycles 2 and 3. The first focus group allowed the researcher to examine the participants' views of different evaluation tools and the format of the toolkit, as well as to develop a better understanding of the users' needs, skills, and ideas about evaluation. Simultaneously, it was used to choose the most suitable research methods for the new evaluation toolkit, with both, the future users of the product and the end users of evaluations. The questions were asked in an interactive group setting where participants were free to talk with other group members. The second focus group was very similar to

the first, but with the aim of assessing GMFRS staff members' views about the newly developed evaluation toolkit.

Gibbs (1997), Sarantakos (2005), and Bryman (2008) note that the role of the moderator is significant in the process as he/she has to be able to manage the group setting effectively. In both occasions the group could freely discuss the given topics, however my role as the researcher and the moderator was to lead the conversation by asking the questions. In the first focus group, the researcher was also seen as a participant in the process, but did not want to influence their opinions by guiding conversations or by expressing personal views. Bryman (2008) discusses the question of control, and debates when the researcher has to take charge of the conversation so that time is not wasted on irrelevant topics. It is important to control the flow of the conversation and ensure each participant has the opportunity to voice their views, but the researcher also found it important to let the participants bring additional topics to the conversations as they were good indicators of their concerns and revealed areas for further research. Focus group was a particularly useful method for this study as it facilitated engagement with multiple stakeholders at the same time. Some of the participants had severe restrictions on their time due to the seniority of their roles in the organisation, and the method allowed effective and time constrained decision making and interaction between the participants.

Gibbs (1997) argues, that the benefits of focus group research include gaining insights into people's shared understandings of everyday life, and the ways in which individuals are influenced by others in a group situation. However, the researcher has to be aware of the group dynamics as it can help identify the conditions that promote interaction and open discussion of participants' views and experiences within groups (McClaran and

Catterall 1997). Even though it is interesting to observe how the participants interact in a group setting, the situation can also severely affect the flow and quality of the conversation. Participants can feel intimated by others' opinions and worry about the impact of their views on their personal career (Sarantakos 2005). It is relatively easy for participants to influence each other, and in the environment this focus group was run, the participants from the lower rank structures could have felt intimidated by the presence of the senior staff members. The questions asked during the sessions were not considered controversial, and the participants had very similar levels of skills and understanding of evaluation which made the situations more equal and less threatening.

5.5 Reflections

All the different research stages of this study contain the researcher's reflections on the success of the research activities, and on personal learning. Action research does not only observe and describe, it encourages the researchers to place themselves in the inquiry (McNiff and Whitehead 2002). One of the central principles of action research is that the researcher learns by reflecting on his/hers own practice (Winter and Munn-Giddings 2002). Reflecting allows the researcher to distant themselves from the research, and review feelings, thoughts, and actions (Rudolph et al. 2001); engaging in inner thoughts and dilemmas helps to plan the next action research Cycle (Marshall 2001). Reflection after the event is helped by careful observation during the event and by good planning before the event (Dick 2002); hence several models of reflection were reviewed. Models developed by Kolb (1984), Gibbs (1988), and Atkins and Murphy (1994) were found to be useful in laying out the different stages of reflections, as shown in Figure 8.
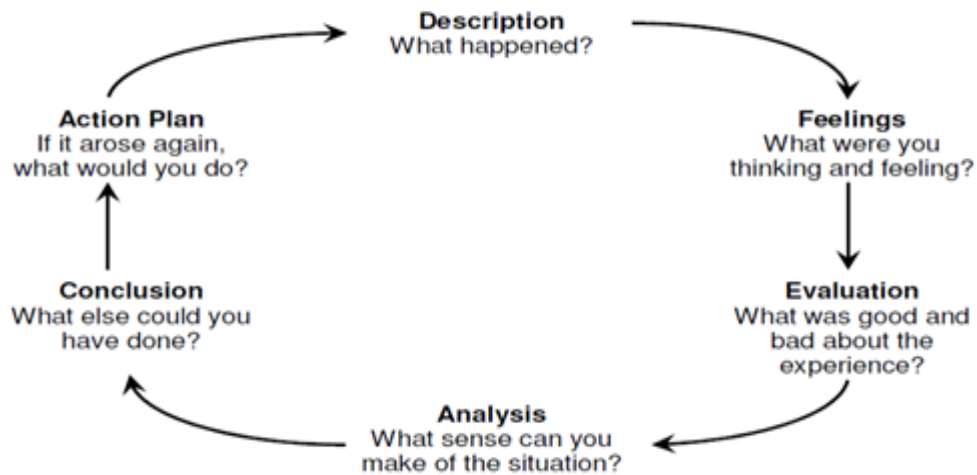
Figure 8: Gibbs' model of reflection (1988)

Boud et al.'s (1985) model of reflection focuses on how the reflective practice produces new perspectives of the experiences, similarly to Borton's (1970) "what - so what – now what" model which encourages the reflector to go beyond the initial experience by adding a further, "now what" dimension. Argyris and Schon (1974) also demonstrate how reflective thinking has to go beyond purely observing outcomes and assessing possible solutions, to assessing the values and assumptions behind these actions. This action research study utilised John's (1994) model of reflection because it provided the most structured approach to reflection. The researcher had limited experience of reflective practice, and a model that provided detailed assistance with the process was needed to add depth to the reflections. The researcher found the following questions set by John (1994) were helpful in the reflection process:

1. Description
   - Write a description of the experience.
   - What are the key issues within this description that I need to pay attention to?

2. Reflection
   - What was I trying to achieve?
   - Why did I act as I did?
   - What are the consequences of my actions? For me; for people I work with.
   - How did I -- and others -- feel about this experience when it was happening?

3. Influencing factors
   - What internal factors influenced my decision-making and actions?
   - What external factors influenced my decision-making and actions?
   - What sources of knowledge did or should have influenced my decision making and actions?

4. Alternative strategies
   - Could I have dealt better with the situation?
   - What other choices did I have?
   - What would be the consequences of these other choices?

5. Learning
   - How can I make sense of this experience in light of past experience and future practice?

6. How do I NOW feel about this experience?
   - Have I taken effective action to support myself and others as a result of this experience?

7. How has this experience changed my way of knowing in practice?

The questions set out by John (1994) were useful in deciding what to focus on when reviewing the research activities and what kind of issues to record. It was helpful to have a set of specific questions to answer because at the beginning the reflective practice did not feel natural, most likely due to lack of experience. The questions assisted the

researcher to view the research activities in a different light, and to think about them from a variety of perspectives. The reflection process was also beneficial in terms of personal learning. Not only did it help to highlight areas that required further study or attention but also taking time out of the research to reflect was useful in sorting out some of the unsolved problems.

5.6 Recording Research Activities

There are two different ways to record qualitative interview data: tape recording and note taking (Dawson 2009). This action research study used the latter because it was most suited to the situation. Robson (2002), Sarantakos (2005) and Bryman 2008 advocate tape recording, (or in the absence of a recorder, a note taker could be used) because poor recording, ways questions are asked, and misunderstandings can lead to unreliable data (Sarantakos 2005). The researcher was aware of the disadvantages of the note taking method, but people may find the recording inhibiting (May 2001, Stringer 2007), and, as discussed in section 4.4 there was a lot of resistance to the research project, maintaining a good working relationship with the GMFRS staff was more important. Even though the researcher had a dual role of an employee and a researcher, the researcher wanted to be "one of them" rather than an external person obtaining information. The note taking technique kept the situation more informal. Everything that the participants said was written down in front of them to maintain transparency, and all the notes taken were always available to the participants. To introduce a quality measure to the interview process, the interview questions and the data were peer reviewed; in the cyclical process of action research, data are gathered and peer reviewed to demonstrate the reality of the practice and its potential impact on others (McNiff, 2000). Prior to the interviews the

questions were pre-tested on members of the project team (four people). The pre-testing was carried out in order to seek out feedback on the questions, the questionnaire layout, and to ensure the questions were easy to understand. The researcher's analysis of the interview data was reviewed by the researcher's PhD supervisor soon after the interviews. It was carried out to see if the supervisor agreed with the researcher's analysis of the data, and if the themes that the researcher had used were representative of the data. The research also utilised the researcher's notes of meetings and observations. In these kinds of situation, tape recording would not have been practical, and the researcher accepts that the notes are subjective – based on the researcher's view of the situation. Copies of the notes from the one-to-one meetings are not included in the thesis because they are considered to be sensitive information. The number of people involved in evaluation activities at the time of study was limited and, as such, this practice guarantees their anonymity.

5.7 Data Analysis

The researcher used two different techniques for data analysis. The interviews were analysed using thematic analysis, and the documents, focus groups, and observations using content analysis. In this section, a description of the techniques is provided, and more detailed information about the analysis of each method is given in the Cycles.

Thematic analysis is one of the most commonly used methods of qualitative analysis, but a "poorly demarcated, rarely-acknowledged, yet widely-used qualitative analytic method" (Braun and Clarke 2006:4). It is very similar to other qualitative analysis techniques, but unlike discourse analysis, conversation analysis and grounded theory, it is not linked to a particular philosophical position (Howitt and Cramer 2008, Braun and

Clarke 2006). Thematic analysis is similar to grounded theory, because the aim is for the researcher to find themes that represent the data (Aronson 1994, Braun and Clarke 2006, Howitt and Cramer 2008). However, thematic analysis, unlike grounded theory, is essentially independent of theory and epistemology, and can be applied across a range of theoretical and epistemological approaches (Braun and Clarke 2006). The analysis follows the steps set out by Howitt and Cramer (2008:333):

1. Code data, i.e. apply brief verbal descriptions to small chunks of data;
2. Alter and modify the analysis in the light of experience and as ideas develop;
3. On the basis of the codings, the researcher then tries to identify themes which integrate substantial sets of these codings;
4. The researcher needs to identify examples of each theme to illustrate what the analysis has achieved; and
5. There is no reason why researchers cannot give numerical indications of the incidence and prevalence of each theme in their data (Howitt and Cramer 2008). Hence analysis can be represented in a qualitative or quantitative manner.

According to Byrne (2001), the researcher decides the process of identifying and grouping the data under the chosen themes in order to communicate the findings simply and efficiently. Even though it is important to be able to communicate research findings effectively, thematic analysis also allowed the researcher to view the participants' answers from different perspectives and handle the data more effectively. It has to be noted that the researcher dealt with analysis question by question, and aimed to establish consensus within each of the questions, rather than across the questionnaire. The aim of the interviews was very precise: to understand barriers to evaluation and to examine the

usability of the toolkit. Interviews can produce vast quantities of fascinating and insightful data hence it can be challenging to stay focused on the actual research questions. The themes gave a good idea of the most common barriers that needed to be overcome in order to embed evaluation in the organisation, as well as how to improve the final version of the evaluation toolkit.

The documents, focus group notes, and observation notes were analysed using content analysis. Prior to a content analysis, the researcher develops a research question, and coding categories, and then progresses to find information regarding them in a text (Robson 2002, Franzosi 2004). It is similar to thematic analysis, but differs in a way that thematic analysis aims to find the themes in the context (Franzosi 2004). The concentration of the document analyses were on the following predetermined themes: drivers for organisational change, reason for community interventions, and the need for evaluation.

5.8 Conclusion

In this chapter all the research methods used in this study: document analysis, focus groups, interviews, and observations were introduced. The aim of this chapter was to provide an overview of the methods, rather than explicit accounts, as they will be discussed in more detail in the Cycle chapters. A justification for utilising John's (1994) reflective model, which was found to be the most comprehensive one of the models reviewed, was also provided, and the chosen note taking technique explained. A discussion about the data analysis techniques, thematic analysis, and content analysis, was also provided.

6. CYCLE 1 – UNDERSTANDING THE ORGANISATION'S EVALUATION NEEDS

In this chapter, details are given of the research activities carried out and the findings made from the first Cycle of the action research study developing evaluation capacity to GMFRS. It was already known that GMFRS had an existing evaluation framework in place which was not being utilised to its full potential. Hence, the aim of Cycle 1 was to understand why this was the case, and to analyse the operating environment in which the new evaluation framework would be embedded. The chapter is divided into five sections. The first section describes the research methods used in the Cycle, and the second summarises the findings made. The third examines the findings in the light of relevant literature, and in the fourth the researcher reflects on the achievements of the Cycle. The final section concludes this part of the study, and summarises actions to be taken in future Cycles.

6.1 Description of Research Activities

The Cycle started with an examination of the aims, objectives, and outcomes of the current community initiatives run by GMFRS to better understand their characteristics. Figure 9 summarises the different stages of Cycle 1.
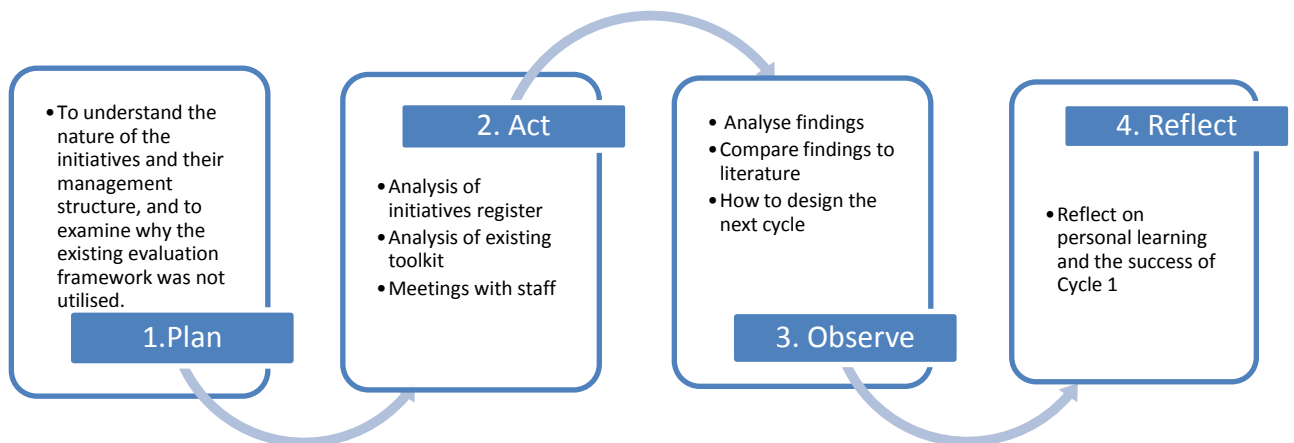
Figure 9: Cycle 1

Analysis of Initiatives Register

GMFRS had a central register in which details of the initiatives that were delivered in the Boroughs were collated. The register was not a monitoring or an accountability tool, but rather a collage of all the initiatives taking place in the Greater Manchester area. The register -- a large data sheet -- was stored in the GMFRS Headquarters, and the department in charge of the register, Central Service Delivery (CSD), relied on the Boroughs to provide all the necessary information. A content analysis, with the aim of assessing the characteristics of the initiatives, and understanding why and how they were run, was undertaken. GMFRS had indicated that they did not have the necessary skills or expertise to evaluate the type of outcomes the initiatives were producing (KTP Grant Application and Proposal Form 2007); hence it was essential to undertake research to understand their nature and to be able to choose the best evaluation methods for their assessment. In the content analysis, every initiative was assessed for evidence about aims, objectives and

outcomes, because clear and explicit project objectives are a precondition for assessing the effectiveness of a service (Phillips et al. 1994).

Analysis of Existing Toolkit

A document analysis of GMFRS procedural guidelines that steer the organisation's initiatives management and evaluation activities was carried out. The documents analysed included the project initiation document, CSD1, and the existing evaluation tool, known as the 'Evaluation Proforma'. The aim of the analysis was to understand why they were not utilised by GMFRS personnel, by concentrating on the following questions:

1. Were the documents easy to use? That is to say, did they include instructions, and sufficient amount of detail about different evaluation tools, techniques and processes?; and

2. Were they aimed at the right level? Bearing in mind the limited evaluation skills of GMFRS personnel, the documents should be designed for a non-experienced user with no knowledge of research and evaluation design.

Meetings with Staff

During the first Cycle, various meetings, related to the delivery of the community initiatives, were organised between the researcher and GMFRS personnel. These included one-to-one meetings, specific to the research project, with senior personnel from the Boroughs and personnel involved in the delivery of the community initiatives, and general committee meetings organised by GMFRS (the researcher was invited as a guest to

observe). The aim of the meetings was to introduce the researcher and the research project to the organisation, and act as an opportunity for the researcher to learn more about the organisational activities, and future users of the new evaluation framework. The meetings took place between April and December 2008 during normal working hours. The selection of participants was based on recommendations given by a senior staff member who was heavily involved in the development of the new evaluation framework. All of the one-to-one meetings took place in the participant's offices or working environments, and the group meetings in the GMFRS Headquarters' meeting rooms. All of the participants, apart from one, were uniformed members of staff, and the rank of the participants varied from a fire fighter to the most senior personnel. In the one-to-one meetings, notes were taken in front of the participants, and in group meetings the researcher's notes were based on observations of group dynamics and the discussion topics. The meeting notes were analysed by using content analysis where the researcher was looking for evidence of GMFRS staff member's evaluation needs.

6.2 Findings

Initiatives Register

The register showed that there were 381 initiatives running between July and September 2008 across Greater Manchester. The headings used in the register to gather information about the initiatives are detailed in Table 4.

Table 4: Initiatives Register

Heading	Details
Date	Start date of the initiative
Department	The name of the Borough where the initiative was running.
Initiative Name	Name of the initiative.
Links to National Outcomes (NO) and LAA Indicators	Boroughs were asked to identify to which national indicators the initiatives contributed to. FRS related indicators: NI 33: Arson incidents, NI 49: Number of primary fires and related fatalities and non-fatal casualties, excluding precautionary checks.
Lead Officer	Person ultimately in charge of the initiative.
Departments	The name of the central department which the initiative was linked to.
Theme	Describes the wider theme of the initiative. Categories included: Partnership Working, Community Fire Safety, Community Engagement, and Offender Rehabilitation.
Sub Theme	A more detailed description of the theme, categories included: Working with Young People, Arson Reduction, Home Fire Risk Assessments, Road Safety.
Equality and Diversity Impact	Low/Medium/High – The level of impact the initiatives have on certain diversity groups.
Diversity Strands	To which diversity strand the initiative contributes (Age/Disability/Gender/Race/Religion or Belief/Sexual Orientation)
Impact Assessment	Boroughs were asked to state if an impact assessment of the initiative had been carried out.
Expected Outcomes	The outcomes the initiative was meant to deliver.
Time Span	Details about the length of the initiative.
Other Comments	Comments about the initiative, its aims or delivery methods.
Evaluation Date	Date when the initiative will be/had been evaluated.
Funding Arrangements	Details of funding bodies/arrangements.

The register collected a vast amount of information about the initiatives, however as shown in Table 4, it did not assemble details about the projects aims, objectives, inputs, or outputs. The register seemed to mainly collect information for external assessment purposes by concentrating on national indicators, diversity strands and equality and diversity impact, but these revealed very little about the initiatives content. The 'Theme' and 'Sub Theme' columns gave some indication of the initiative's goals by vaguely describing the target audience and the type of initiative: Community Fire Safety, Offender Rehabilitation, Working with Young People, Arson Reduction, Home Fire Risk Assessments, and Road Safety. In Figure 10, the percentage of initiatives that had provided information about the main headings, listed in Table 4, is represented. The most frequently completed columns in the register were: 'Date', 'Department', 'Initiative Name', 'Departments', 'Theme', 'Sub Theme', 'Equality and Diversity Strands', 'Expected Outcomes, 'Time Span', and 'Other Comments'.
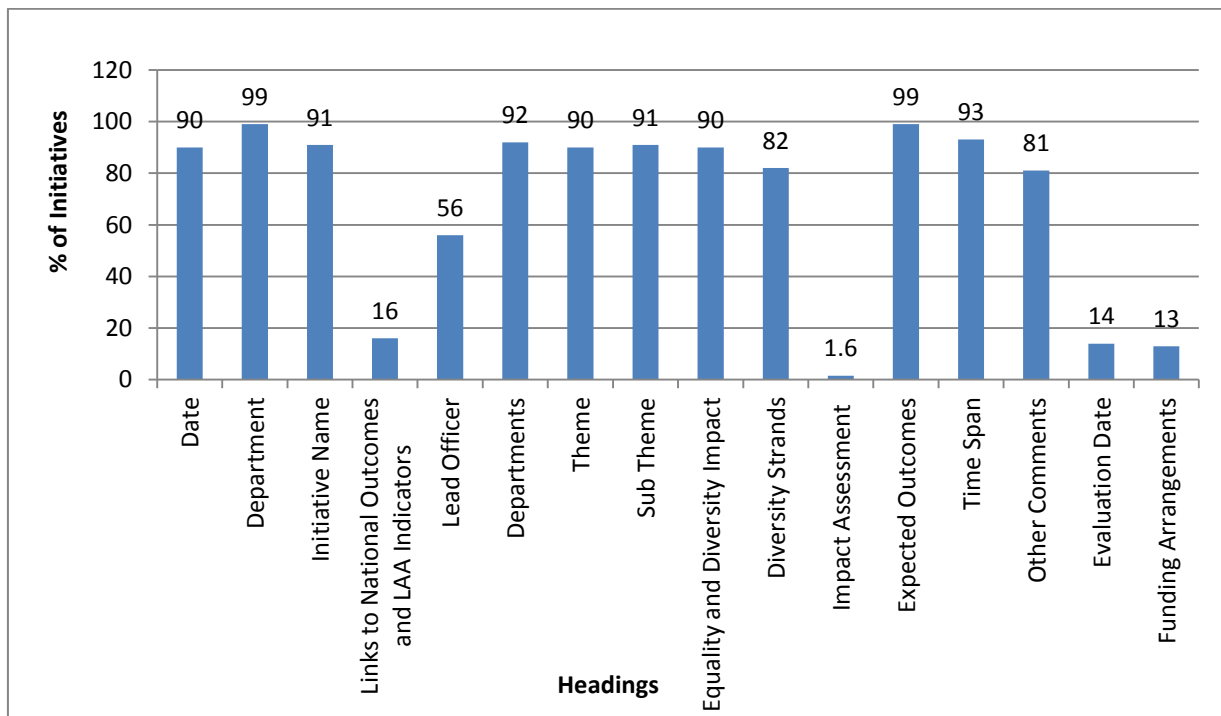
Figure 10: Percentage of initiatives provided information (n=381)

The initiatives are principally driven by National Performance Indicators (KTP Grant Application and Proposal Form 2007); however, as shown in Figure 10, only 16% of the initiatives had details of the Indicators to which they were linked. During meetings with GMFRS staff members, it was indicated that partnerships/externally funded initiatives were more likely to be evaluated. However, when trying to examine if any links between external funding and evaluations could be established, it was noted that only 13% of the initiatives had any details of their funding arrangements, and only 14% of initiatives had been evaluated, or were planning to be evaluated. Of those initiatives that had provided funding details, 6% were evaluated, and even though they were all partnerships, the evidence was not strong enough to create links between external funding and evaluations. In Figure 11, the funding sources of the initiatives that had provided information about funding have been explained.
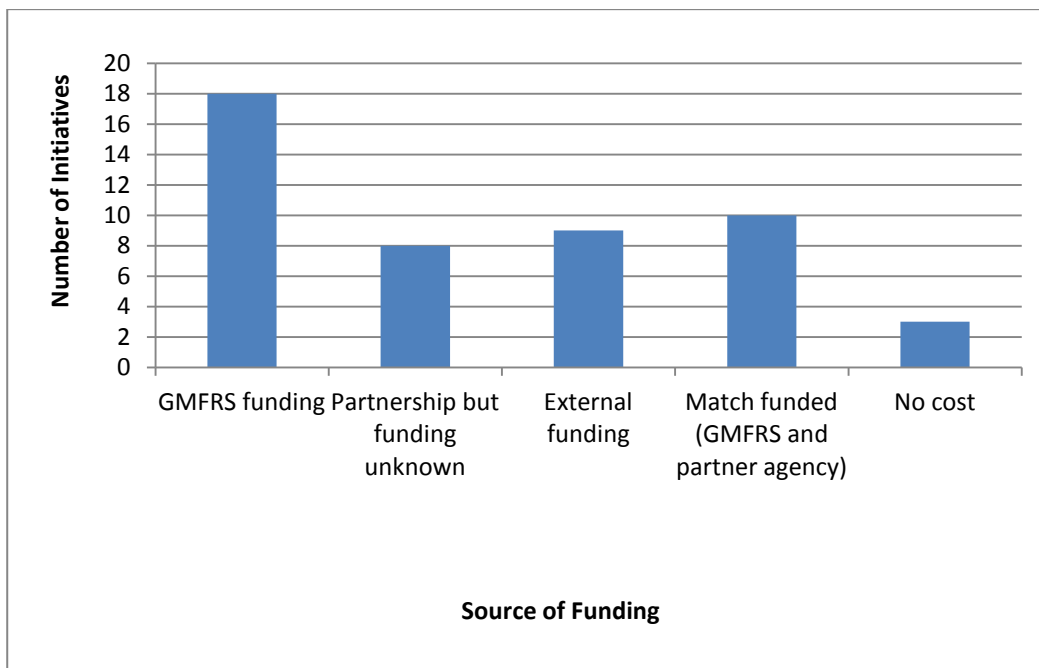
Figure 11: Funding sources of initiatives (n= 48)

It was assumed that the most informative section for understanding the ultimate goals of the initiatives would have been the 'Intended Outcomes' column, and for the majority of the initiatives, (99%), information had been supplied. However, a content analysis of the column revealed that the organisation did not have a common understanding of the term because the information provided in this column varied noticeably, as shown below:

Reduce the number of domestic dwelling fires

Reduce the fire of crime including fire related crimes.

Increase awareness of personal and community fire safety in the home.

Reduction in the number of alcohol related accidental dwelling fires and subsequently the number of injuries and deaths.

500+ premises currently without smoke detection being provided with smoke detection.

3000+ premises occupied by renting pensioners will have HFRA's carried out.

Issue of 6,000 tea-light holders to people identified as using tea-lights unsafely.

Delivery of 4000 'portable ashtrays'.

For the use of the HVP outside of GMF&RS County.

Standard procedure for all Ops crews to adopt incidents.

Hard wired smoke alarms, Streetclean initiatives. Potential to pilot domestic sprinklers.

Improvement in the quality of CFS activities undertaken.

Increased HFRA's in the Indian Community

Raise awareness of dangers of fire to school children in the Borough.

Locally run youth engagement programme, 6 week duration, run by operational fire fighters.

Education on the importance of smoke alarm ownership and fire escape plans.

Reduce the number of drug/alcohol related fire deaths and injuries.

Reduction in road traffic collisions.

3200 HFRA's per year undertaken by volunteers.

As shown in the statements above, the content of the 'Intended Outcomes' column was very varied, and gave little insight about the actual outcomes of the initiatives. Therefore, to better explain and understand the content, the researcher used the following categories: outcome, output, description and unkown/?, to classify the above mentioned statements[2]. In Figure 12 a diagrammatic analysis of the categories is provided.

Outcome = goals, stated as specific changes in participants' behaviours, knowledge, skills, status, or functions (W.K. Kellogg Foundation 2001:6):

Reduce the number of domestic dwelling fires

Reduce the fire of crime including fire related crimes.

Increase awareness of personal and community fire safety in the home.

Reduction in the number of alcohol related accidental dwelling fires and subsequently the number of injuries and deaths.

Reduction in road traffic collisions.

Reduce the number of drug/alcohol related fire deaths and injuries.

Improvement in the quality of CFS activities undertaken.

Output = the direct products you anticipate from activities. These may include the types, levels, and targets of services delivered through your intervention/activities. (W.K. Kellogg Foundation 2001:6):

---

[2] The statements are reproduced under the categories to demonstrate how they were classified.

500+ premises currently without smoke detection being provided with smoke detection.

3000+ premises occupied by renting pensioners will have HFRA's carried out.

Issue of 6,000 tea-light holders to people identified as using tea-lights unsafely.

Delivery of 4000 'portable ashtrays'.

3200 HFRA's per year undertaken by volunteers.

Increased HFRA's in the Indian Community.

Description = an explanation of initiative that does not include details of outcomes or inputs:

To have young members of the community assist in the delivery of organisational objectives whilst providing them with a valuable insight into the Fire Service and the world of work.

Partnership working to work towards making Bury smoke free. We are now trying to share date with this agency. The primary care trust has the data as to which households will have smokers.

Once process is in established identify suitable staff within social services to undertake HFRA'S on our behalf.

Locally run youth engagement programme, 6 week duration, run by operational fire fighters.

Education on the importance of smoke alarm ownership and fire escape plans.

Raise awareness of dangers of fire to school children in the Borough.

? = Information that could not be categorised:

For the use of the HVP outside of GMF&RS County.

Standard procedure for all Ops crews to adopt incidents.

Hard wired smoke alarms, Streetclean initiatives. Potential to pilot domestic sprinklers.
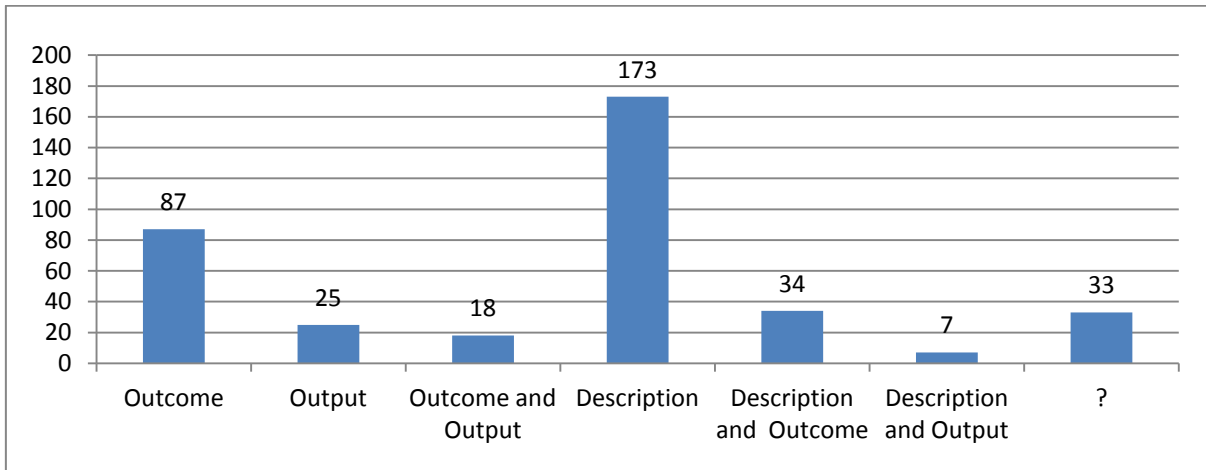
Figure 12: Analysis of 'Outcomes' column (n=377)

As shown in Figure 12, only 23% of the information in the 'Intended Outcomes' column could be defined as outcomes. Nearly half of the content (46%) was descriptions of the initiatives activities, and the rest (31%) were mixtures between descriptions and outputs, descriptions and outcomes, and information that did not fall into any of these categories. This was a clear indication that GMFRS would have to have commonly agreed definitions for all project related terms, and that the initiatives needed better defined goals.

Existing Evaluation Methodology

The organisation had two documents that Boroughs/project initiators could use to record initiative activities. The first one was the project initiation form, CSD1, which had to be submitted to the centrally managed Partnerships and Innovations Management (PIM)

board[3] to gain approval to start a project. The PIM board acted as a project management/scrutiny structure, and invited staff members responsible for the delivery of the initiative to report on the success of the initiative on a regular basis.

In the CSD1 form, the following headings were used to record information:

1. Title and Brief Description of Borough Proposal;

2. Purpose of Borough Proposal (e.g. nature of partnership initiative etc. that the policy/procedure is intended to support);

3. Corporate Aims and Objectives (Insert the Service or Borough Plan references(s) that the proposal is designed to support);

4. Performance Management (Insert which BVPIs/LPIs should be improved by the introduction of the proposal); and

5. Implications (Tick the relevant boxes below to indicate which functional areas may be affected by the proposal. Specify which section(s) within each area may be affected & provide details as appropriate.)

At the beginning of this action research project GMFRS had indicated that they had no methodology/knowledge to measure the outcomes of the initiatives (KTP Grant Application and Proposal Form 2007). However, as the CSD1 headings show, GMFRS was not encouraging staff to think about and communicate the aims, objectives, and outcomes of the initiatives. The headings focused on establishing the impact of the initiative for the organisation: its purpose, how it supported corporate aims, how performance would be managed and what functional areas would be affected. None of the

---

[3] This was only applicable to initiatives where the total cost would exceed £5000. Initiatives costing less than £5000 were managed by Boroughs, with no input from the Headquarters.

headings encouraged staff members to describe why the initiative was run, how it was run, what it aimed to achieve, in what timeframe or with what intended results.

The PIM board encouraged GMFRS staff to use the existing evaluation tool, the Evaluation Proforma, to evaluate project achievements. The tool had been developed as a result of an externally delivered evaluation training course organised for senior management. An examination of the proforma revealed that it did not include any guidance on different evaluation methodologies or processes, or any details of evaluation tools that could be used to assess the success of the initiatives.

The evaluation proforma encouraged the users to record the following types of information:

1. Description of Initiative (Initiative/Project Aims and Objectives)
2. Project Restraints
3. Type of Evaluation
4. Outcomes and Impact of Project
5. Cost Effectiveness
6. Community Reaction
7. Partners Involvement/Reaction
8. Sustainability
9. Learning and Recommendations

Each section included a brief description of the kind of information the author was expected to provide. The proforma was not an evaluation tool, but a template for reporting

back on evaluation findings. The meetings with staff members had already revealed that the proforma was not very widely used because the staff felt it was not fit for purpose. Two members that had attended the evaluation training course mentioned that the training (and the proforma) were only given to senior members of staff who were not responsible for evaluations, and the skills and knowledge were never passed down to those who would have required them. At the time, the organisation did not have the knowledge or resources to provide evaluation guidance to anyone wishing to use the evaluation proforma.

<u>Meetings</u>

Conversations with various staff members showed that there seemed to be a lack of clarity about the different evaluation processes, in particular the following issues:

- The distinction between traditional social science research and evaluation. These two activities were considered synonymous;

- What aspects of the project to evaluate and when. Some believed that evaluation was only used to measure the longer term impact of projects, such as social, behavioural and economic change;

- Unfamiliarity with evaluation methods. Some had been able to use quantitative research methods to analyse project's outputs, but had not felt comfortable measuring the softer outcomes that most of the initiatives produced. Some had, unsuccessfully, tried to apply the same quantitative methods to measuring the

softer non-tangible outcomes, which had led to the belief that soft outcomes were impossible to evaluate; and

- Evaluation activities were not carried out in any logical order. Most of the evaluation processes took place at the end of the project, including evaluation planning, data collection and reporting.

6.3 Review

The first task of the Cycle was to assess the portfolio of all the community initiatives so that their characteristics and diversity could be ascertained. This was followed by an examination of the existing evaluation tool, and meetings with GMFRS staff to gain a better understanding of the organisation's evaluation needs. The key findings of these activities were that the initiatives lacked clearly defined aims and objectives, GMFRS staff were not encouraged to communicate the projects goals, and the existing evaluation materials were not detailed enough for a non-experienced evaluator. There was also confusion about the role of evaluation, and what kind of processes it entailed.

These findings were interesting, as in Chapter 3, 'Contextualising Study', the term evaluation was examined, and its role established. Evaluation, according to Rossi and Freeman (1993:5) is "the systematic application of social research procedures assessing the conceptualisation, design, implementation, and utility of social intervention programmes". The main purpose of evaluation is to improve programme delivery and make it more responsive to client needs (Rutman and Mowbray 1983). The benefits of evaluation

include increased knowledge, improved decision making, and evidence of changes/success/impact/efficiency/effectiveness/overall quality/areas for improvement (Posavac and Carey 1985, Patton 1997, Robson 2000, Boulmetis and Dutwin 2005, Davidson 2005). It has also been argued that key evaluation activities included "looking at objectives" and "judging the worth of projects" (Scriven 1991, Rossi and Freeman 1993). The GMFRS initiatives register was used to collect information about various aspects of the initiatives, but it did not encourage the Boroughs to identify the project's aims and objectives. The lack of understanding about the project's goals had important implications for evaluation, as clear and explicit project objectives are a precondition for assessing the effectiveness of a service (Phillips et al. 1994). Professionals and managers need to have clear ideas of what the programmes are setting out to achieve. Without such clarity it is unlikely the goals and intended purpose of service will be achieved (Phillips et al. 1994:69). Evaluation looks at the achievements of a project, and how and why these have occurred, and clear, specific and measurable programme goals are part of good evaluation conditions (Patton 1997). As the projects did not have clear direction, it became questionable how to collect meaningful data in order to assess the true impact and achievements of the activities. Hence, it was decided that to create a successful evaluation framework, GMFRS should be encouraged to articulate the project achievements in a clear manner. The study of the initiative management structure and the existing toolkit also highlighted the need for improved evaluation materials, including guidance on how to evaluate and how to use research methods in evaluation. This also confirmed there was a need to construct common understanding of the meaning of evaluation, what it could achieve and how it was done. The initiatives also needed clearer goals in order to carry out evaluations in a systematic manner, which meant that common definitions for the key words -- aims, objectives, outputs, and outcomes -- would have to be created.

6.4 Reflections

Reflection

The aim of the first Cycle was to understand the organisation's evaluation needs. The research activities chosen to do this were document analyses and meetings with GMFRS staff. The methods were chosen because of their neutral nature; document analysis was thought to be fairly unobtrusive, and meetings were a good way to create contacts within the organisation. With regards to the research activities, the best possible strategies for the situation were employed because, in addition to trying to address the research objectives, there was also an attempt to build the organisation's trust in the researcher, and gain support for the research project. In terms of alternative strategies, focus groups or other types of group consultations could have been used, and they would have probably been a more efficient way of collecting data during a time restricted research project. However they could have affected the relationship building exercise at this very fragile stage.

This first phase was very challenging, because the researcher was not well known to the organisation, and was asking staff members to provide information that was not readily available. The key themes that shadowed the research activities were lack of contact and lack of data. The researcher was not based at the GMFRS for the first four months, and had great difficulty locating the right people and data sources. It was also realised that the organisation had not prepared the staff for my arrival, only two people were aware of my existence. There was some confusion over the researcher's role - some thought the researcher's purpose was solely to evaluate projects, and they were very keen

to give her all evaluation related tasks, but very disappointed when the wishes could not be fulfilled. As some staff members were not aware of whom the researcher was and why she was collecting sometimes very sensitive information, they tried to avoid any contact the researcher tried to make. Many of the contact attempts and information/meeting requests were ignored. Reflecting on this, it could have been due to two factors: 1) the dispersed structure of the organisation, the Boroughs relied on different communication methods to the Headquarters, and 2) most of the key personnel were based at the GMFRS Headquarters. In order to deliver a successful project the researcher asked to permanently be moved to the GMFRS premises, and sit in a centrally located office. A communications plan, that utilised all the possible communications channels in the organisation, was also designed to create awareness of the research activities. These steps exposed the research to the organisation to such extent that the researcher started receiving invitations to meetings and contacts from the Boroughs.

The findings of the research activities, especially of the analysis of the initiatives register, were unexpected. It was surprising to note three issues: 1) a very large public sector organisation did not use a coherent project management methodology to design or manage the delivery of the initiatives, 2) as the aim of evaluation is to assess the achievements of a project, no links had been made between lack of project planning and the poor quality of existing evaluations, and 3) so little data existed about the projects. Also a great number of the case studies in the evaluation and ECB literature (for example Pawson and Tilley 1997, Minnett 1999, Stevenson et al. 2002, Davidson 2005, Monroe et al. 2005, Cohen 2006, and Forss et al. 2006) concentrate on evaluation techniques, producing quality evaluation, and on the utilisation of evaluation findings. It was surprising to note that none of them had examined, or even considered, barriers to internal

evaluation practices or the impact that lack of project planning and management can have on evaluation activities.

The most influential sources of knowledge were the Kellogg Foundation Evaluation Handbook (1998) that provided a framework to evaluate initiatives that were lacking clearly defined goals, and Patton's Utilisation Focused Evaluation (1997). Patton (1997) examines political and other external factors that influence evaluations and utilisation of evaluation findings, and it was a relief to find some commonalities between the challenges and being able to relate to them. For example, for some time GMFRS had been trying to evaluate the initiatives without a correct evaluation methodology. The process they were using was not giving them the information required, hence the findings were never utilised. Patton (1997) talks a lot about organisations that carry out meaningless evaluations to give the impression of efficient service, and challenges everyone to stop this wasteful practise, and only evaluate if the findings can be used either to improve the project or inform stakeholders.

In terms of success, the ultimate goal for the Cycle was to understand the organisation evaluation needs. The document analyses produced findings that will help to change the evaluation practices of the organisation. More one-to-one contact with the organisation would have been beneficial, as the researcher started to question if there were other barriers to the conduct of evaluations, than lack of evaluation materials and guidance, and inadequately defined initiatives.

<u>Learning</u>

It was very surprising to note how reluctant the organisation was to deal with an external researcher, and it really affected my confidence as a researcher and a research project manager. However, this experience taught me that timing of communications and relationship building are vital for the success of a research project. In the future attention has to be paid to the following issues:

- When joining an organisation ensure the researcher/ the researchers intentions/ the research project are introduced prior to the start of the project, and again directly after the start of the project;

- Create communication channels, with the help of the organisation, at the start of the project;

- Use various communications channels to ensure everyone is reached, and be open about the intentions of the project to avoid misunderstandings; and

- Negotiate contact with right people prior to the project start date.

In terms of the literature – the researcher was trying to understand the nature of the initiatives and why were they not evaluated. Prior to starting the research activities, it was thought GMFRS purely did not have the skills to evaluate and required more training about evaluation. But the study of the initiatives register revealed that the problems were a lot more complex. The findings led the researcher to examine project management and its

links to evaluation, which the researcher would have never done without this experience – there is surprisingly little about this in the literature. These research exercises presented evaluation in a new light, not as an isolated activity, but as a process interlinked to effective project management.

6.5 Conclusion and Future Actions

This chapter contains details of all the research activities carried out during Cycle 1 and the findings made. The Cycle consisted of two different types of research activities: document analyses and several meetings with GMFRS staff. The activities of this Cycle were carried out for two reasons, to understand the operating environment in which evaluation would be embedded, and to examine barriers to the use and embedment of evaluation into GMFRS. The key findings of the Cycle were: a) initiatives lacked direction, and b) the existing evaluation tool was not detailed enough for the use of personnel with no previous evaluation experience.

These findings have highlighted the need for the following future actions, to be addressed in Cycles 2 and 3:

1. The organisational understanding of evaluation – clarity about the role of evaluation, and what it can achieve;

2. The importance of project management – embed a project planning model which can underpin evaluation activities;

3. Evaluation materials – A 'how to' guide that would guide the user through the different stages of evaluation, and the use of research methods. Define key concepts; and

4. Gain a deeper understanding of barriers to evaluation.

7. CYCLE 2 - DEVELOPING EVALUATION MATERIALS

In this chapter details are given of the research activities carried out, and the findings made from the second Cycle of the action research study developing evaluation capacity to GMFRS. The first Cycle identified that GMFRS's community initiatives lacked direction, and the existing evaluation tool was not detailed enough to be used by personnel with no previous evaluation experience. The aim of the second Cycle was to develop a new evaluation toolkit, and assess its usability amongst non-experienced personnel through three trial evaluations, and analyse the barriers within GMFRS to the internal evaluation of community initiatives by interviewing staff involved in evaluation activities. The chapter is divided into five sections. The first describes the research methods used in the Cycle, the second summarises the findings made. The third examines the findings in the light of relevant literature, and in the fourth the researcher reflects on the achievements of the Cycle. The final section concludes this part of the study, and summarises actions to be taken in future Cycles.

7.1 Description of Research Activities

This Cycle started with an examination of other evaluation toolkits to identify best practice in the field, and this was followed by a focus group that chose the research methods for inclusion in the GMFRS evaluation toolkit. A prototype evaluation toolkit was developed as a result of the toolkit reviews and the focus group, and it was then tested on three different community initiatives. The Cycle also included interviews with GMFRS staff members to learn more about barriers to internal evaluation practices. Figure 13 summarises the different stages of Cycle 2.
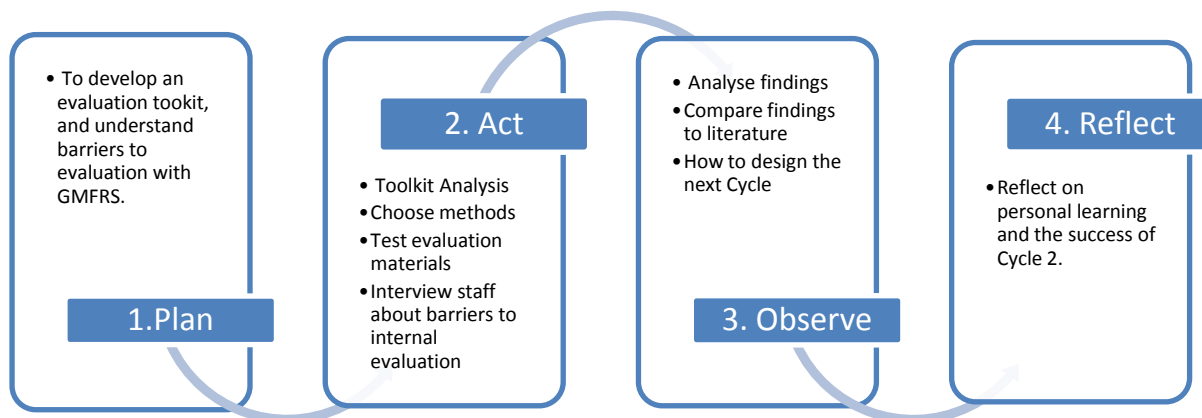
Figure 13: Cycle 2

Toolkit Analysis

The document analysis from the first Cycle revealed that GMFRS needed improved evaluation materials that would: a) clearly define all the key words associated with evaluation, and b) guide the user through the different stages of evaluation and the use of research methods. GMFRS had expressed the need for the following type of evaluation toolkit (KTP Grant Application and Proposal Form 2007):

- An evaluation toolkit and decision making framework to guide and support GMFRS users in designing and undertaking evaluations. To include: evaluation tools and techniques, information on parameters of application and validation protocols.

- GMFRS personnel have very limited evaluation skills, and no expertise in how to focus on measuring evaluative outcomes, hence, the toolkit has to be sensitive

to the variability and complexity of GMFRS public protection initiatives, and robust and capable of being utilised by inexperienced personnel.

The project core team[4] had also indicated that the toolkit would have to be easy to read, short in length, and available electronically and as a paper copy. Apart from these fairly general requests, the organisation had not given any clear direction of the format of the toolkit; therefore it was decided to carry out an analysis of other available evaluation toolkits. The researcher decided to build a prototype toolkit based on the requests of GMFRS, and the best practice from other toolkits, then seek feedback on its usability and test its functionality with GMFRS staff.

The following documents were included in the analysis:

1. Aim Higher West Yorkshire, 'Evaluation Toolkit' (n.d.)

2. Aim Higher Greater Merseyside, 'Evaluation Toolkit' (2006)

3. Sport England, 'Monitoring and Evaluation Toolkit' (n.d.)

4. Centre for Health Promotion, 'Evaluating Comprehensive Workplace Health Promotion Info-Pack' (2005)

5. Evaluation Assistance Center, 'Evaluation Handbook' (1995)

6. Institute for Work and Health, 'Guide to Evaluating the Effectiveness of Strategies for Preventing Work Injuries' (2001)

7. Department of Health and Human Services, 'How to Evaluate Safety and Health Changes in the Workplace' (2004)

8. J. Harvey, 'Evaluation Cookbook' (1998)

---

[4] See p.63 for further information.

9. Annabel Jackson Associates, 'Evaluation Toolkit for the Voluntary and Community Arts' (2004)

10. W.K. Kellogg Foundation, 'Evaluation Handbook' (2004)

11. The World Bank, 'Monitoring and Evaluation: Some tools, methods and approaches' (2004)

All the toolkits, apart from 5, 6, 8 and 10 were selected from an internet search, in which the researcher was seeking toolkits that were easily accessible for everyone looking for evaluation guidance. The following search words were entered into an internet search engine (Google[5]): 'evaluation toolkit', 'evaluation framework', 'evaluation non experienced user', and 'evaluation guidance'. The other toolkits were recommended by two academics from the University of Salford. The toolkit search and review was carried out between January and March 2009, and the aim of the analysis was to identify best practice in the field, and to find designs and ideas to fulfil the following quality criteria, that were composed from the findings of the first Cycle and the requests GMFRS had made:

1. Tools/methods to be based on sound research and analytical methods supported by existing research;

2. Methods must help the evaluators reach their conclusions on the basis of the evidence gathered;

3. Toolkit must be suitable for use by (trained) GMFRS staff who are not evaluation specialists;

---

[5] Google was the only search engine used.

4. Toolkit must be capable of being imparted via a training course, delivered by internal GMFRS personnel;

5. Toolkit must permit analysis of intended and unintended outcomes; and

6. Toolkit must include written support, guidance and examples for the evaluators.

Choosing Methods

GMFRS had indicated that the staff members had very limited research skills, and in the first Cycle it was identified that the existing toolkit did not include any guidance on using research methods for evaluation purposes. Meetings with GMFRS personnel from Cycle 1 also revealed that unfamiliarity with evaluation methods had led to the belief that some of the initiatives were impossible to evaluate. The researcher decided to produce a list of the most commonly used research methods, and to hold a focus group where GMFRS staff members could choose the methods to be included in the toolkit. The researcher reviewed different research methods and soft outcome measurement tools, and produced the following list of methods that could be used without excessive training (can also be found in Appendix B):

- Focus Groups
- Nominal Group Technique
- Interviews
- Questionnaires
- Surveys
- Daily Diary or Personal Journal
- Content / Document Analysis
- Observations
- Goal Setting and Presentation of Material in Portfolio
- Tests and Scales
- Cost Effectiveness

The selection was based on the methods used in the previously reviewed evaluation toolkits and the works of Patton (1997), Dewson et al. (2000), Robson (2002), and Bryman (2008). During the focus group the researcher explained to the participants how the methods were used, and the advantages and disadvantages of each method. The participants were then given an opportunity to discuss the methods, and choose the ones they thought to be best suited for GMFRS users. The role of the researcher was to manage the meeting, and also act as participant in the conversations, but not in the decision making.

The focus group was held on the 27th February 2009, during normal working hours in the GMFRS Headquarters' meeting room, and was attended by ten participants, in addition to the researcher. The attendees were selected after consultation with two GMFRS senior managers, and eight of them were uniformed members of staff. Seven of the participants, according to the senior managers, represented other senior staff members from Boroughs and the Headquarters, responsible for the future evaluation activities of the organisation. One of the participants was a lower rank staff member who had been actively involved in (voluntary) evaluations and was invited because of his background and interest in research. Two participants were academics from the University of Salford, and attended the meeting to provide expert advice on the use of different research methods.

Notes were taken in two ways: a GMFRS administrator was asked to take minutes of the meeting, and record the decisions made. These notes were circulated to all attendees after the meeting. The researcher also observed the situation[6], and kept personal notes of

---

[6] Participant observation

the participants' views and behaviour. The decisions made during the focus group were used as the analysis of this research activity.

Testing the Prototype Toolkit

After a prototype evaluation toolkit that addressed all the requirements (more detailed account of the toolkit will be provided in section 7.2) had been developed, the researcher decided to seek feedback from GMFRS staff members on its practicality, and test it on three community initiatives to realise its usability. The previous toolkit was not fit for purpose because it did not include enough detail about the different evaluation processes and data collection methods, hence this feedback and testing phase was crucial to ensure the new toolkit was suitable for the organisational needs and users' skills. Initial feedback was collected from the "core team", described in Chapter 2: two senior uniformed officers, directly and non-directly in charge of community initiatives, one non-uniformed member of staff in charge of some of the GMFRS research and evaluation activities, and an academic from the University of Salford. These people were chosen because they were the close to the project, and in positions where they had impact on the future success of the toolkit. The participants were given copies of the prototype toolkit, and after they had reviewed the content, one-to-one meetings were organised to discuss the participants' views. The meetings took place between May and July 2009, during normal working hours in the GMFRS Headquarters, in both the participants' offices and meeting rooms. Notes were taken by the researcher in front of the participants. The feedback was used to improve the prototype product.

After the initial improvements to the toolkit (discussed in more detail in section 7.2), the chosen evaluation approaches and methods were tested on three community initiatives. The researcher contacted three Borough Commanders, chosen by a senior GMFRS officer, to identify what initiatives were running at that time of the year. The Borough Commanders requested the toolkit to be trialled on the following three initiatives:

1. Moss Side Fire Station Boxing Club (trial carried out between June-October 2009)

   The local fire fighters had experienced a negative attitude from some of the young people residing in the area, and recognised that much of the work they have to do is a direct result of the youths' actions. They also discovered that boxing had a certain kudos in the area, and it could increase the respect that an individual acquires from their peer group, hence they decided to set up a boxing club to extend their influence into the local youth groups. Boxing as a sport mirrors the personal attributes fire fighters value; both require courage, discipline, dedication, focus and respect for others. It is a sport that anyone can participate in regardless of race, creed, gender, or social and economic status. The club has been running on a voluntary basis since June 2008, and runs 15 sessions for four different groups (youth, disabled, women, and mixed adult) every week. In addition to teaching boxing skills, the youth classes are designed to educate the gym users about respectful and disciplined behaviour as well as the difference between fighting and boxing (Evaluation Summary Moss Side Fire Station Boxing Club November 2009).

2. "Reach for the Sky" Reading Scheme (trial carried out during August - October 2009)

St George's School is one of the two primary schools on the Hag Fold estate, Atherton. According to the project documentation, this estate has higher than average levels of crime/anti-social behaviour and was the subject of one of four Community Improvement Projects delivered in 2007- 08 lead by the FRS in the Wigan Borough. It was recognised by St George's Primary School that some children do not have access to reading material or support from parents in the home environment to develop their reading skills. The school has utilised the support of lay adults from its church to mentor children in reading, and recognised the benefits of mentoring being delivered by people who are seen by the wider community as positive role models. GMFRS was approached by the Deputy Head of the School with a request to utilise fire fighters as these positive role models (Evaluation Summary "Reach for the Sky" Reading Scheme Wigan 2010).).

3. Firefly (trial carried out during September - December 2009)

The Firefly aims to engage with young people aged 11 – 17 years. The accredited five-day course is based at a working fire station and run by fire fighters. It offers young people an energetic physical challenge and the chance to learn skills in fire fighting, home safety awareness, first aid, team building, communication and personal development. The target youths may have low self-esteem or confidence, and as a result they may be displaying anti-social

behaviour including deliberate fire setting and malicious calls to the emergency services. The programme is also encouraging young people who have offended, or may be at risk of offending, to take part (Evaluation Summary Firefly Manchester North March 2010).

The trials ran from April 2009 until December 2009. All of them were conducted by GMFRS staff and supervised by the researcher. The focus was on assessing the training needs of GMFRS, and observing the GMFRS staff members' ability to carry out evaluations. It was voluntary for the people running the initiatives (community volunteers and GMFRS staff members) to take part in the evaluation trials, and everyone involved in the initiatives participated. After the initial contact with the participants, they received training and guidance on the evaluation procedures, and help from the researcher in designing evaluation plans and materials. The initial plan was to give a copy of the toolkit to the project managers, who, with the assistance of his/her team and researcher, would then carry out the evaluation as set out in the toolkit. However as all of the participants were new to evaluation, and did not have to the necessary resources, mainly time, to do this on their own, a decision was made that they would test out the evaluation materials and processes set out in the toolkit (logic model, evaluation plan and research tools) with the help of the researcher. First the researcher talked all the participants through the toolkit, and asked them to use the logic model to communicate the project's achievements so that an evaluation plan could be drafted. The evaluation plans were drafted jointly, and included details about the type and the timing of the evaluation, as well as the aspects of the project that would be evaluated and the data collection methods that would be used to answer the evaluation questions. This was followed up by some research methods training, which in the case of Moss Side Fire Station Boxing Club was provided by an academic

from the University of Salford who specialises in community interviewing. The staff members who participated in the other two evaluation projects received written and verbal interview training from the researcher; these were two were small scale projects with few research participants. Reading Scheme had four data collectors and Firefly two in addition to the researcher.

Participant observations were carried out by the researcher during the evaluation activities. During the observations the researcher focused on the following:

- How participants were able to complete logic models, evaluation plans, and use data collection tools, such as questionnaires. Were they able to undertake these activities on their own or did they require assistance and, if the latter, what kind of assistance was required?
- Were they able to competently undertake data collection? Were these additional activities able to fit into their daily activities?
- What would they do with the evaluation findings – utilise or ignore?

The researcher kept written records of the observations throughout the evaluation trials.

Interviews

To help understand existing evaluation practices in GMFRS, and barriers and constraints to the utilisation of the existing evaluation toolkit, a series of one-to-one, semi-structured interviews with GMFRS staff members involved in the delivery of the initiatives were conducted. The interviewees were selected from a list of 39 GMFRS staff

members who had been trained to use the existing evaluation toolkit. After a discussion about the trained personnel with a senior manager from GMFRS, it was established that four of them had left the organisation, and 13 were employed in roles where they would not have used the existing toolkit or have an involvement in the organisations' future evaluation activities. From the remaining 22, seven could not be contacted[7]; the interview invitations were sent to 15 people, from which 12 accepted to participate.

The interviewees were all uniformed members of staff in senior positions either in the Boroughs or the GMFRS Headquarters. The seniority of their position meant that they would not be directly responsible for carrying out evaluations but, rather, in charge of commissioning evaluation activities to lower rank staff members. The interviews were carried out during normal working hours between September and October 2009 in the interviewees' offices, and conducted by the same interviewer. At the beginning of the interviews, the participants were given a copy of the interview questions (Appendix C), and the questions were also asked by the interviewer. Everything that the participants said was written down in front of them to maintain transparency. This note taking technique was used to prevent taking any further time from the participants; however they were informed that if they wanted to see any of the material it would always be available to them.

The interviews were analysed by using thematic analysis, and the common themes in the answers were then used as the findings of the interviews. The aim of the interviews was to examine the views of the participants in relation to the existing toolkit, the existing

---

[7] The contact details provided were either out of date, or the participants did not reply to the researchers contact.

evaluation practices of GMFRS, barriers to internal evaluation practice, and ideas about future evaluation activities of the organisations.

7.2 Findings

<u>Toolkit Analysis</u>

With regards to the quality standards, detailed in the previous section, none of the toolkits fulfilled all of the criteria. The toolkits analysed could be divided into three categories:

a) Emphasis on evaluation processes, very little or no research method guidance;

b) Emphasis on research methods, very little or no evaluation process guidance; and

c) Complete guide that covers both evaluation processes and research methods (In depth, lengthy high level document, suitable for large scale projects or policy evaluations).

Table 5 illustrates the categorisation of each toolkit reviewed.

Table 5: Toolkit Categories

Toolkit	A (Emphasis on evaluation processes)	B (Emphasis on research methods)	C (Complete guide)
1. Aim Higher West Yorkshire	✓		
2. Aim Higher Greater Merseyside		✓	
3. Sport England		✓	
4. Centre for Health Promotion			✓
5. Evaluation Assistance Center		✓	
6. Institute for Work and Health		✓	
7. Department of Health and Human Services	✓		
8. J. Harvey		✓	
9. Annabel Jackson Associates			✓
10. W.K. Kellogg Foundation			✓
11. The World Bank		✓	

The majority of the toolkits came under category B, and concentrated on providing guidance on social science research. The researcher learnt from these toolkits that it was important to provide the following information for the GMFRS personnel:

- Clear guidance on how to design and use research methods;

- Examples and ready to use templates; and

- Guidance on qualitative and quantitative data analysis.

Toolkits that came under category A, where the emphasis was on evaluation processes, demonstrated the importance of:

- Clear structure and guidance on how to use it;

- Define all key words;

- Introduction to evaluation: what it is, what it can do and how it is used;

- Practical and easy to follow guide to the different stages of evaluation; and

- Evaluation planning templates, and examples of completed forms.

Category C toolkits were complete evaluation guides of very high quality and demonstrated that:

- GMFRS toolkit should not be too long;

- GMFRS toolkit does not have to cover every aspect of evaluation, such as history of evaluation;

- Include a reading list for those who wish to read more about evaluation;

- GMFRS toolkit has to be written in an easy to understand and follow manner; and

- Use logic models to help evaluators to clarify project goals.

Focus Group

The focus group was held so that GMFRS staff members could choose research methods, suitable for the organisational needs, to be included in the toolkit. The researcher

presented a list of methods to the participants, and in Table 6 the research methods chosen by GMFRS are shown:

Table 6: Research Methods

Research Method	To be included in the Toolkit
Focus Group	✓ (with a manual)
Nominal Group Technique	✓
Interviews	✓
Document Analysis	✓
Questionnaires and Surveys	✓ (Only questionnaires)
Diaries	
Observations	
Goal Setting and Presenting Material in Portfolio	
Tests and Scales	✓ (with further research)
Cost Analysis	

The methods were chosen unanimously, and the group felt the chosen ones were easy for everyone to use with some further guidance. Questionnaires were chosen but surveys, defined by the group as large-scale mass questionnaires, were believed to benefit from the input of a more experienced evaluator, and would not be suitable for the toolkit. The group also discussed the use of tests and scales to measure soft outcomes such as changes in behaviour, confidence and attitudes, and it was decided that more research about the different tests would have to be carried out to identify the best ones for GMFRS.

Many of the tests needed the user to have a background in psychology, and a good understanding of research practices; hence, it was felt they were too complex for the new toolkit at this stage. The methods that were not chosen to be included in the toolkit at this stage were diaries and observations. It was felt that, due to the subjective nature of these methods, GMFRS personnel might find them difficult to use and analyse. Goal setting was also thought to be too complicated and time consuming. The group decided that cost analysis was not required at the time.

The researcher observed a very important discussion during the focus group. One of the members could not understand how the chosen methods could be used to evaluate GMFRS outcomes. When asked what kind of outcomes the staff member meant, the participant described long term outcomes and national performance indicators such as changes in socio-economic conditions. The person held the same belief that was observed during Cycle 1 meetings, that some of the GMFRS initiatives' outcomes were impossible to evaluate. This lead to a discussion, within the focus group, about a project's short, medium and long term outcomes, the importance of clearly defined and SMART[8] aims and objectives, and what aspects of the project to evaluate and when. It was important to have the discussion during the focus group for two reasons. Firstly, most of the attendees were senior officers, and if they could get the evaluation processes right, they could then pass on the information to people working with them, and secondly, this highlighted the need to develop an evaluation toolkit that clearly defined when the different stages of evaluation should be carried out.

---

[8] Specific, Measurable, Achievable, Realistic, Timed

<u>Testing the Prototype Toolkit</u>

The new toolkit (Appendix D), developed after the toolkit review and the focus group, provides a step by step guide to evaluation, and is designed to be used by both non-experienced staff members as well as personnel who have previous evaluation experience. The toolkit addresses three different types of evaluations: process, outcome, and impact, and provides tools to measure both hard and soft outcomes. The toolkit is divided into four sections, as described below:

1. Understanding Evaluation

   The first part of the toolkit provides an introduction to evaluation, and explains to the reader why, when, and what to evaluate. Those familiar with the different stages of evaluation are given the choice to use the evaluation summary at the back of document as a checklist to ensure they have considered all the necessary aspects of evaluation.

2. Planning Your Evaluation

   The second part of the toolkit concentrates on planning an evaluation. It guides users through all the different stages of evaluation planning, and includes examples of completed evaluation plans for those who are new to evaluation.

3. Gathering Evaluation Evidence and Making Sense of Findings

   The third part is a guide to the data collection tools that can be used as part of any evaluation process. The tools were chosen from a range of research methods in a focus group attended by various GMFRS personnel. This section

provides guidance on how to use the methods and analyse the findings. Also, information about resource requirements, and advantages and disadvantages of each method is provided to allow the users to choose tools based on their needs and experiences.

4. Reporting and Sharing Evaluation Findings

The final part of the document explains how the users should report and share their evaluation findings. There is also a section on recommended reading for those who wish to learn more about the different evaluation and/or research method techniques, and the appendix section provides the users templates for evaluation planning, and a list of key word definitions.

The initial feedback from the three GMFRS staff members and an academic from University of Salford revealed that the prototype toolkit was easy to read and an appropriate length. The following suggestions for improvements were made:

- Guide and a template to measure cost effectiveness;

- Expand on data analysis;

- Guidance on choosing assessment indicators;

- An evaluation report form that prompts people to record unanticipated outcomes (whether negative or positive);

- More detailed examples of different methods; and

- Some minor layout changes.

The evaluation trials showed that the participating staff members were not able to use the prototype toolkit on their own, and needed guidance with all the different stages of evaluation. They were able to carry out some of the data collection activities without an input from the researcher, but the planning and design stage proved very challenging, hence, the researcher had to have more input in the trials than initially planned. The trials also showed that initiatives did not have clearly defined aims and objectives, but the use of logic models helped staff to clarify the initiatives' goals. These were important findings since the aim of the trials was to test the toolkit approach, including how staff members could use the toolkit, how they would find the time to carry our evaluations, and what kind of support they would need in the future. The findings were used to form the basis of a recommendation paper that would be taken to the management to inform them of the staff members' evaluation needs.

The observations made during the trials could be categorised under the following themes:

- Support and Guidance: GMFRS personnel were very supportive during the evaluation planning process, however, perhaps due to the limited use of evaluation within the Service, not all had a clear understanding of the different stages of evaluation and the design of data collection tools. For example, staff found it challenging to draft logic models that clarified the goals of the projects, and evaluation plans. They had a very good idea of what they wanted to achieve with the projects but the projects were lacking clearly defined goals (Appendix E). This was observed in all the three trials, and could have been due to lack of training and unfamiliarity with the evaluation tools the researcher was asking the

participants to use. However, it was most likely linked to what was observed in Cycle 1, most of the GMFRS initiatives lacked clearly defined aims and objectives, because the organisation did not encourage staff members to identify detailed goals for the initiatives.

- Management/Supervision: The trials showed that all evaluation activities had to be managed very closely, and that a lack of supervision would have resulted in incompletion of the tasks. When staff were not sure how to do something, instead of asking advice, they preferred to ignore the task. This became evident in the data collection stage, and the researcher had to have constant contact with the participants and manage them through the tasks;

- Resource Use: Evaluation is very resource intensive, and time is the biggest commitment an evaluator has to give to the project. The researcher faced a lot of difficulties during the data collection stages as some staff members were not prepared for the time commitment; hence the evaluation timetables had to be altered multiple times (Appendix F); and

- Benefits: The evaluations allowed a closer investigation and comparison of the intended and actual aims and objectives, and the delivery methods of the projects. This resulted in redefining the project's goals to make them more realistic and achievable. The trial participants also realised the data collection tools could be used in every day monitoring of the initiatives, to ensure internal and external quality standards are met.

When the interviewees were asked "What do you think is the value/purpose of evaluation?" all of them were able to identify at least one key aspect of evaluation:
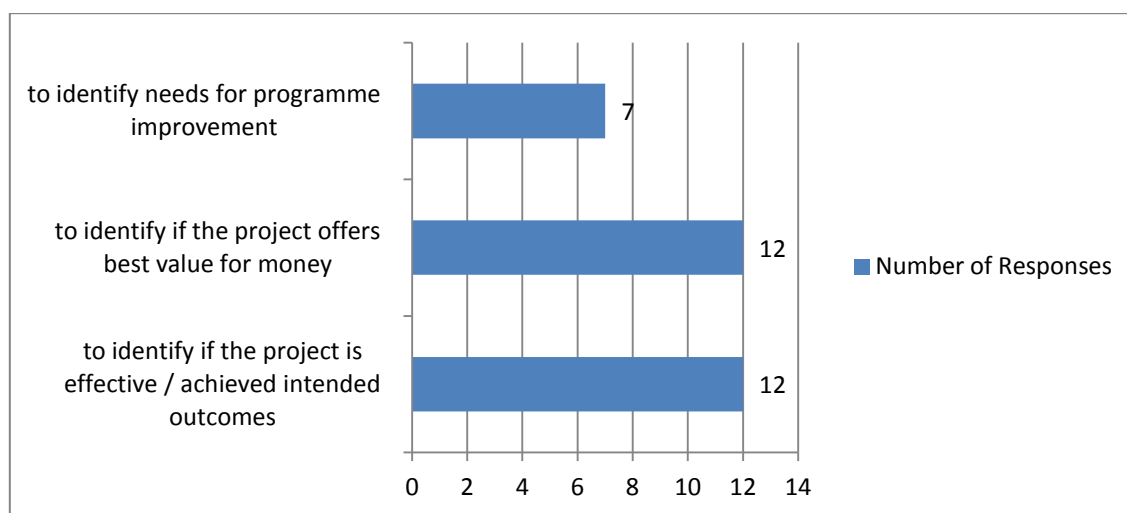
Figure 14: Value and purpose of initiatives (n=12)

Participants had a very clear understanding of the ultimate goal of evaluation:

> "If you do it correctly, evaluation will determine if you're achieving your results and shows what to do with the project - continue or cut short. It also helps to find out what delivery methods have worked and why."
> *Participant 2*

> "Value and purpose of evaluation is to reflect on initial aims and objectives and to identify how over a period of time the initiatives have developed and changed. Making yourself to ask the questions if it still achieving the aims and objectives, and have the aims and objectives changed and whether they need to change. Ultimately to identify if project has achieved its outcomes and at what cost in terms of finances and resources."
> *Participant 8*

Participants also demonstrated how they had considered the benefits of evaluation by discussing the importance of identifying if the initiatives were delivering value for money:

> "For FRS essential to carry out something robust and valuable to ensure initiative is effective and economical. Need to make sure its achieving its objectives within given resources (time and money). Need to do full evaluations and take other political issues of the area into account; evaluations cannot be done in isolation."
> *Participant 10*

> "Primarily to ensure initiatives we embark on meet the intended outcomes in terms of organisational goals. We need to know what we want to achieve and how we have achieved it. Value for money. We tend to move on the next initiative without thinking could it have been done better."
> *Participant 12*

However, five out of the 12 respondents agreed that, in general, GMFRS personnel do not have a very clear understanding of the purpose and activities of evaluation. And seven believed that only some members of the organisation, mainly senior managers and those who had attended evaluation training, understood the different processes and purpose of evaluation.

Some participants believed this was due to the lack of organisation wide awareness and training:

> "Not at all apart from the 42 people who took part in the evaluation training. People don't have any idea of what evaluation is."
> *Participant 1*

> "Not at all levels. All of those who attended the evaluation training course know what evaluation is all about, but I'm not sure if they would be able to evaluate."

Participant 2

"No. Limited amount of people understand key evaluation processes and use them. [They] don't understand purpose of evaluation because they haven't been explained to, and therefore don't understand the value of evaluation (and don't do it)."
Participant 9

However others believed it was more to do with the culture of the organisation and

the working habits of the individuals:

"No they are trained to get the job done and move on. Don't stop and think how things could have been done better. Even though debriefs have just been employed they are a level of evaluation, as they give a chance to talk about the experience. But they aren't using the skills on the delivery of the initiatives."
Participant 3

Not all personnel. Limited amount of knowledge and understanding at this level and not fully embedded into our role. Majority of initiatives watch based and they should know principles of project management and evaluation. It is a new world for watch officers."
Participant 8

The interviewees were also concerned about the quality of existing evaluations.

When they were asked: "Do you think the evaluations that are currently undertaken by the

GMFRS are of value to the organisation?", seven out of the 12 thought the quality of the

current evaluations hindered their value, and five of those thought, the biggest threat to the

quality was the person who carried out the evaluation: the same person who manages the

project also evaluates the project – which, according to the interviewees, leads to biased

results. Other issues the interviewees were concerned about were evaluation design, the

organisational awareness of different evaluation processes, and poor planning, all of which

meant evaluations were not capturing the right kind of information.

Some believed the quality issue was caused by the lack of interest in the task, and they

showed some serious concern over this:

> "They are not taken seriously. Evaluations are written by sponsors or project managers and they only represent bias evidence. I don't know who even looks at them."
> *Participant 1*

> "Not always, because question about independence. And problem I have is that some evaluations are done by the project managers and it is very difficult for a project manager to criticise their own projects. I have seen very extreme examples of that – not very successful projects have been made into brilliant ones by the project managers."
> *Participant 4*

Some were unsure of the quality of the existing evaluations because they felt the

existing evaluations did not focus on the assessing the right aspects of the

initiatives, or were not carried out in the right order:

> "[Evaluations are] not capturing soft outcomes, not using range of evaluation techniques that are available, and they are not planned properly. Evaluation should be determined at the beginning, not at the end."
> *Participant 2*

> "I think – I would say yes but its work in progress. The biggest failing is that evaluations are only done at the end. Now beginning to understand the need to start at the beginning and carry on during project life."
> *Participant 5*

> "Are of value, but hard to think how they are of value. One value is reporting to many arenas externally. Value is limited as the evaluations are not as in-depth and infinite as they could be. Same applies to initiatives. And therefore we can't give true feedback and we may stop them unfairly because of poor evaluations."
> *Participant 9*

"A number are robust and credible, and carried out in methodological manner. However, a number of evaluations are written like the project has contributed to too many outcomes, and it is very unlikely that they have done that. I&P initiatives tend to be evaluated like this, and all of them use different methods, so it's difficult to evaluate the evaluations."
Participant 10

On a positive note, one respondent felt that evaluations had been very beneficial to

the initiatives, and had improved efficiency and effectiveness:

"Definitely to initiatives because by undertaking evaluations we have improved effectiveness and efficiency. And they have resulted in changes in directions. Not 100% sure as an organisation that we are obtaining all the benefits of evaluation at the moment. Think we are to use it to share good practice and identify good value for money. But cannot see any intangible benefits to support the initiatives that have been identified as issues in the evaluation process."
Participant 8

All of the interviewees knew there was an evaluation toolkit available to use at

GMFRS, but only five had used it. And of those:

- three found it repetitive;

- one thought the level was too high; and

- one thought it was a template that does not take triangulation into

  consideration.

A participant who had used it believed that even though it was not fit for purpose, it was

better than nothing:

"Yes – I use it all the time. It is slightly repetitive but better than nothing. Other people don't probably use it because the level of the toolkit is too high. Some people say they don't use it because they haven't been trained to use it. And I think there is also confusion over role – no one knows whose

role is to evaluate. Unless evaluation is brought as part of project management it won't get done. Everyone also needs to understand that evaluation won't work at the end, it needs to be built into the project management process."
Participant 1

And, even though the rest found it laborious and repetitive, no one completely dismissed it:

"Yes, and I have used it. The existing toolkit is template and doesn't take triangulation into consideration. Have had no difficulties using it, but only because I know what I'm doing, have been on the evaluation course."
Participant 2

"Yes, and have used it. I thought there were unnecessary duplications. It was repetitive, asked same things again and again; and laborious."
Participant 5

"Yes I know it exist and yes I have used it. It has gone through number of changes. I have been on a training course and the GMFRS developed its own proforma, which is repetitive and don't cover some of the key issues GMFRS has to deal with. The toolkit doesn't give anything to evaluate the direct benefits to community and social capital you get from initiatives. But the toolkit is a good starting point."
Participant 8

The rest of the interviewees (7) had never used it because they were not working on projects that needed to be evaluated. A majority of the respondents (10) felt GMFRS did not provide sufficient resources to support the planning and completion of evaluations, and of those:

- five felt there were no evaluation tools;

- four felt the level of training was inadequate;

- four thought evaluation should be incorporated into the project management framework; and

- three felt there was insufficient level of support and guidance to carry out evaluations.

The lack of resources, and how it affected the use of evaluation in the organisation, was highlighted in many of the participants' views:

> "No – it is a process that is not bolted into something. Officers are not given tools to evaluate and audit their activities as part of their training. Evaluation is not explicitly talked about, other than hard outcomes."
> *Participant 3*

> "Resources are limited. Limited training. Person leading the Beat Sweep evaluation attended the evaluation course and has collected some documents but doesn't feel he's competent to fully evaluate it. There is no comprehensive toolkit or tools, only a proforma which is not suitable to evaluate all projects. Measures outputs but can't evaluate the soft outcomes. And haven't got the skills - don't know how to do it really. Can tell something reduced by X% but can't tell why. Can't measure full value of activities, just go off gut feeling."
> *Participant 5*

But also a real concern, in addition to the lack of resource, was the lack of time to carry out evaluations:

> "I suppose it doesn't, but this is not major criticism. Early stages of embedding culture of evaluation/understanding of evaluation need to be front loaded with resources. But have put people like me on training course and give toolkit without any assistance with that. They assumed that I would embed the culture into the Borough. I have given my experience to other people in Borough and now they face the same problems."
> *Participant 8*

> "No. People see evaluation as an extra workload and always feel it's something that's done at the end. People don't think it has to be done

throughout the project – mind set. We don't have different evaluation methods for small and large projects, only one method (existing toolkit). For small projects the method is too much. Existing toolkit put people off – too much to do if only a small project. Projects that go through PIM board have more outcomes to evaluate and can use the existing form. Problem is the evaluation info is not used in anyway."
Participant 12

A majority of those interviewed (11 out of 12) also believed there were barriers and constraints to evaluation. Looking at the 11 in more detail, this is what they believed to be the barriers to GMFRS evaluation practices:
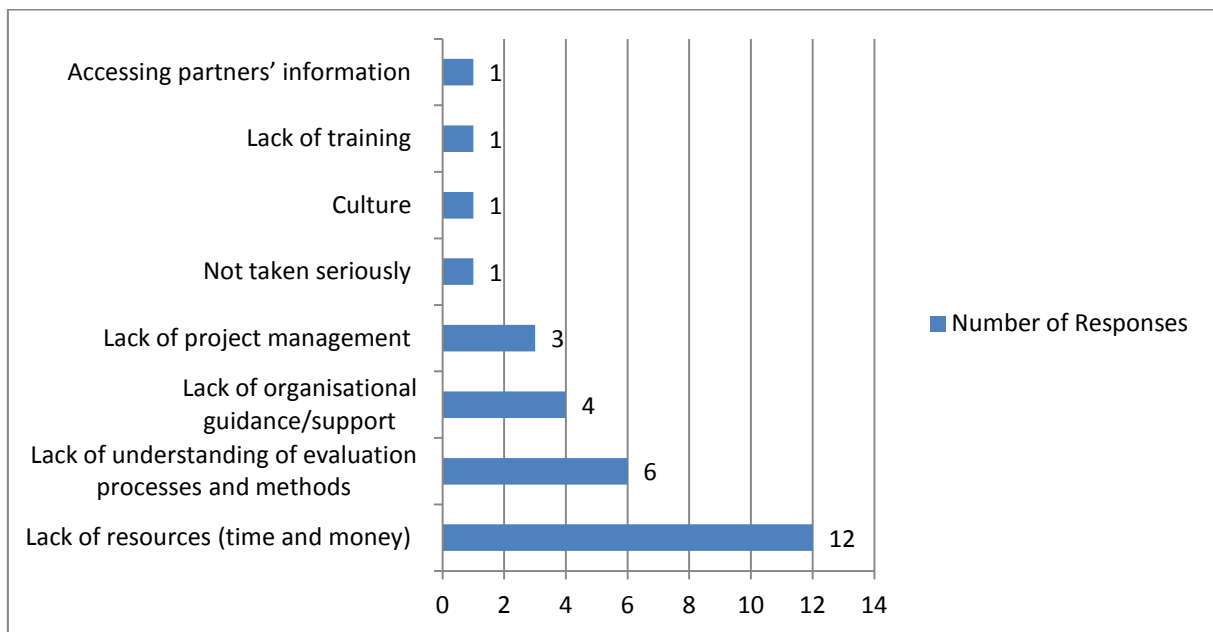
Figure 15: Barriers to Evaluation (n=12)

The answers to this question were very similar to the answers of the previous one; the participants found time, lack of resources and money as barriers to evaluation:

"Biggest restrictions - no clear mandate, no time, no one takes evaluation seriously (even if it is a big and expensive project). Evaluation is not built into the project management, which is a problem. There is no monitoring and midterm reporting structure. Lack of project management means the

cost of the initiatives is too high, and GMFRS needs to stop some of them. There is no way to improve the projects because no one is evaluating them. The projects could be a lot smarter in delivery if they were evaluated."
Participant 1

"Economic constraints due to the economic climate. Cannot spend too much money on evaluation. Independent evaluators are very expensive. Management also has workload and prioritising limitations."
Participant 2

"Resources: time, money and budget available. Without these it's difficult to get full evaluations. Could do half a day evaluation but would it have any value? Support from other people – because they are busy they cannot provide support."
Participant 4

"Yes, resources and time. And if people think they need to do it, to tick a box or don't believe in it, it won't be done properly. Lack of understanding and lack of understanding of methods."
Participant 9

"Biggest one is time, because we don't consider it early enough and are under pressure to do everything at the end. Resources – when going to do it. Finance – external people are very expensive."
Participant 11

From this data it is clear that the interviewees recognised that there are barriers to evaluation, but also seemed to appreciate the usefulness of evaluation activities. They felt positive about monitoring and assessing the initiatives, and thought evaluation was a valuable activity as it helped to address issues related to the effectiveness of the project, identify best value for money activities, and highlight possible future changes to programme delivery. The majority of the participants (11 out of 12) also believed that GMFRS should carry on evaluating the community initiatives themselves:

"If GMFRS is a professional organisation, there should be no problems evaluating the initiatives. Internal audits are already taking place, so the skills should be transferable to evaluation."
Participant 2

"Yes it is not just absence of training but about cultural awareness of the need to evaluate. More about recognising the need for qualitative evaluation of own activities. It's a thought process – evaluation should be part of the culture. Fire fighters don't need another process but rather a change in mind sets."
Participant 3

"Yes and it should be built into some of the standard training courses like step up training etc. Partnerships are now an everyday activity."
Participant 5

When GMFRS personnel were asked about how they would like to see the evaluation processes develop in the future, these were the key themes that emerged:
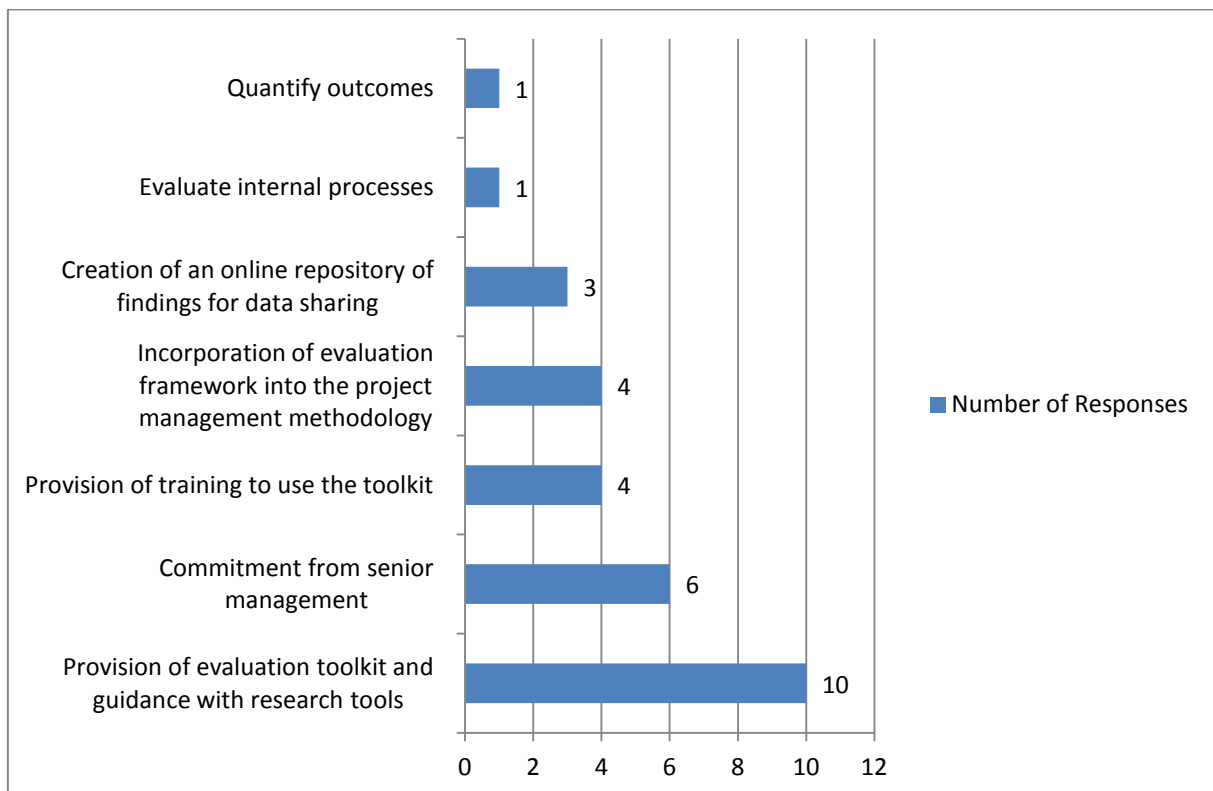

Figure 16: How to Develop Evaluation Processes in the Future (n=12)

A few of the participants focused on the assistance individual members of staff needed to carry out evaluations:

> "Training should be at step up level. Now changes taking place in recruitment structure, as they aim to get more non-uniformed personnel to replace the uniformed, so evaluation training should be part of the new recruitment training. Needs huge commitment from senior management, no commitment at the moment. Commitment is most important as recommendations from lower ranks are not otherwise taken seriously."
> *Participant 1*

> "More training. Real objective evaluations to be done - need a department to do it, but without taking the autonomy away from BCs. Need to do centralised evaluations. Get register on share point. Confirm with BCs that evaluations need to be done throughout the project, not at the end. If SARA model was used we wouldn't enter all partnerships we do now."
> *Participant 7*

Some participants emphasised the need for the organisation to take the lead on evaluation activities, and show intent and support by developing a policy and making the role of evaluation clearer in the organisation:

> "Criteria for evaluation – explanation of what kind of evaluation and resources each type of project needs. Repository of evaluations to share data."
> *Participant 2*

> "Strategic statement of intent – needs to start from up – a policy on evaluation to determine what to evaluate, when and how. Determine methodology for evaluation. Put in place training – and roll out. Have it as a measure in performance management terms and commission external evaluators (evidence lead solutions have done beat sweeps)."
> *Participant 6*

> "Policy and structured approach that's proportional and graduated and directive. Simple and easy to use toolkit. Has to be fully circular system, not just good feedback about the initiative, but need to also show if it didn't work."
> *Participant 9*

"Clear policy. Clear commitment (policy from strategic level). Guidance. Plan for it. Training. Implement process. Audit and review."
Participant 10

7.3 Review

The focus group, evaluation trials, and interviews revealed that the majority of participants did not have a good understanding of evaluation processes. The trial stage and the interviews also confirmed the findings made during the first Cycle that GMFRS staff needed assistance with clarifying project aims and objectives. The trials proved how labour intensive evaluation activities are, and how important it is to monitor them closely otherwise they would not be carried out. During the interviews it was identified that GMFRS staff members found the knowledge gap, and lack of resources and organisational guidance were barriers to carrying out evaluations. The interviewees believed that the organisation would benefit, in addition to new evaluation materials, from further guidance with evaluation processes, and an organisational commitment to evaluation activities. Staff also thought it would be useful to incorporate evaluation into a project management methodology, and create avenues for sharing evaluation results. The researcher's experiences from the focus group, the trial stage, and from the first Cycle confirmed these views.

The research findings of Cycle 2 highlighted that the challenges of this action research project, developing evaluation capacity into GMFRS, were related to a knowledge gap, lack of resource, and lack of organisational guidance. It was known to the researcher that some research had been carried out about the barriers to utilisation of evaluation

findings (Leviton and Hughes 1981, Cousins and Leithwood 1986, Patton 1997), barriers to external evaluators/evaluations (Taut and Alkin 2003, Taut and Brauns 2003), and the barriers to learning from evaluations (Taut 2007, Torres and Preskill 1999). In Chapter 2 of this thesis, 'Background to the Study', it emerged that none of the existing ECB efforts had discussed any challenges in developing the systems, but the results could be compared to Gibbs et al. (2002), Milstein et al. (2002) and Taut and Alkin's (2003) previous work on barriers to evaluation. Taut and Alkin (2003) examined barriers to external evaluation implementation in the context of a university outreach programme, and noted that the staff found the following factors as obstacles to evaluations:

- Human factors - refer to the user's knowledge about evaluation and the creditability of the evaluator. Taut and Alkin (2003) have identified that the less people understand the evaluation activities, the less likely they are to feel positive about evaluation;

- Evaluation factor - looks at how evaluation is conducted – the quality of design, data collection and information; and

- Context – in which the programme exists: political and organisational influences and barriers.

In Taut and Alkin's (2003) study, participants identified human factors as the biggest barriers to evaluation, evaluation factor came second, and the context factor received least attention. An analysis using this framework of the GMFRS barriers is given in Table 7: the human and context factors were clearly identified as the key obstacles.

Table 7: GMFRS Barriers to Evaluation[9]

Human	Evaluation	Context
• Lack of understanding of evaluation processes and methods (6) • Lack of project management (3) • Not taken seriously (1)		• Lack of resources (time and money) (14) • Lack of organisational guidance/support (4) • Culture (1) • Accessing partners' information (1) • Lack of training (1)

Gibbs et al. (2002) have carried out a study about programme managers beliefs and attitudes to evaluation, and identified four factors influencing evaluation behaviour among community based organisations:

- Funding agency expectations;

- Resources (staff time, access to external consultants, funding for operational costs, and computer hardware and software);

- Leadership; and

- Evaluation tools and technology.

Milstein et al. (2002:42) also found resources as one of the main barriers to evaluation, as well as righteous attitudes, which assume that a programme works in the absence of evidence, and resistance that stems from the human instinct to avoid criticism, judgment, and change, and low confidence in the methods of evaluation science. Using Taut and Alkin's (2003) framework, Gibbs et al.'s (2002) findings fell under evaluation and context factors, context being the more influential factor, and Milstein et al.'s (2002)

---

[9] Number in brackets represents the number of participants who mentioned these barriers.

under the human, evaluation and context factors. All of these studies had been carried out in learning organisation with existing evaluation system, and external evaluation assistance readily available. GMFRS experiences of barriers to internal evaluation were very similar to the other three case studies, but differed in a sense that GMFRS staff did not find the evaluation factors as barriers. However, when the organisation's evaluation activities become more established the evaluation factors could become an obstacle. This is an interesting finding, as it can help the organisation to prepare for future issues.

7.4 Reflections – Cycle 2

<u>Reflections</u>

The second Cycle involved three key activities: developing an evaluation toolkit suitable for use by non-experienced personnel, involving GMFRS in the process, and gaining an understanding of barriers to evaluation practices within GMFRS. The focus group was a very successful method. The group was very keen to take part in the toolkit development, and it gave GMFRS staff some ownership of the product. The first Cycle was overshadowed by contact difficulties, therefore discussions, that took place during the focus group, about evaluation outcomes, and how other departments should be included in the toolkit development felt like a breakthrough. For the first time in the project's lifecycle, GMFRS showed support for the research. In Cycle 1, it was realised that GMFRS evaluations were not carried out in any logical order and staff members were not sure what aspects of the project to evaluate and when. Some believed that evaluation was only used to measure the longer term impact of projects, such as social, behavioural and

economical change. In the focus group, the academics and the researcher had the opportunity to address at senior level the false beliefs GMFRS held of evaluation.

The quality of the evaluation toolkits was surprising. Vedung (2010) argues that faith in scientific evaluation eroded in the early 1970s, and there is very little evidence in the evaluation literature of anyone using the scientific methods anymore, however two of the biggest toolkits had chosen to advocate this route. Authors such as Posavac and Carey (1985), Robson (2000) and Boulmetis and Dutwin (2005), argue how evaluation has to be differentiated from social science research, but six of the eleven toolkits treated evaluation as a pure research activity. It seemed that toolkits that had been developed in academic institutions employed this more scientific research based approach, whereas evaluation consultancies preferred a more holistic concept. In Cycle 1, the close relationship between evaluation and project management was investigated, and it was surprising to note that many of the toolkits had ignored this. The only exception was the W.K. Kellogg Foundation toolkit that was directed by an empowerment evaluation approach, which clearly guided the users through the different evaluation activities, and made suggestions on how to time them to correspond to the different stages of a project life cycle. However, reflecting on the content of the toolkits, the differences could be explained by looking at the context in which they were developed, as it was identified in Chapter 3 that evaluation capacity building efforts are context dependent and based on organisational needs. The toolkit developer's background could also be a contributing factor. Alkin (2004) has examined the development of different evaluation orientation, and concluded that the discipline resembles a tree that is built on the dual foundation of accountability and systematic social inquiry. The tree itself is divided into three branches: use, methods, and valuing -- representing the aspects of evaluation the authors' approaches originate.

The second stage of Cycle 2, the trials, demonstrated that GMFRS staff members needed more guidance with evaluation than just the toolkit. It was very surprising to realise how little interest some of the staff members had in the delivery of the projects and the evaluation process. To carry out the interviews concurrently with the trials was a very good decision. They really helped to deepen the relationship with the organisation; however it would have been helpful to carry them out earlier. The information obtained would have been really helpful during the first Cycle, and the one-to-one contact with the staff would have helped to gain more support for the project.

Overall, Cycle 2 was very successful in achieving the planned goals. The trials and the interviews really made the organisation realise what evaluation was and how effective it can be. The staff that participated in the trials had held a fairly negative view of evaluation, but as soon as they realised it was not about assessing their personal work input, but identifying if the projects were achieving their outcomes, the attitudes changed. The whole organisation seemed to become a lot more supportive after they had read the evaluation reports. Reflecting on this, it was probably, because during the first stage no tangible outcomes were produced for the organisation; whereas during this Cycle GMFRS had the opportunity to view the newly developed evaluation materials, and the evaluation trial reports, which made them realise what this action research project could achieve.

Learning

- GMFRS is traditional hierarchy, and it seemed the staff were very keen to participate in the research, as long as they would not have to physically do anything. In the future it is essential to think more about the organisational

culture before deciding how to engage with the organisation. Organisational /
senior management support has to be gained right at the start of the project to
make the research activities run more efficiently;

- Toolkit: it is difficult to decide how much information to include, and how to
  balance the information needs of the different users. All the information was
  new to them, so it was challenging to find the balance between not enough and
  too much detail. But the feedback confirmed it was the right level at the time. It
  is important to include staff in the development process; and

- A good way to get people to improve processes / products is to provide them
  with opportunities to reflect on them. The interviews gave the participants an
  opportunity to reflect on the organisation's evaluation practices, and they had
  surprisingly many ideas on how to improve them.

7.5 Conclusion and Future Actions

This chapter contains details of all the research activities carried out during Cycle 2
and the findings made. The Cycle consisted of four different types of research activities:
document analyses, focus group, observations and interviews. The activities of this Cycle
were carried out for two reasons, to develop an evaluation toolkit for the assessment of the
GMFRS community initiatives, and to examine barriers to the use and embedment of
evaluation into GMFRS. The key findings of the Cycle were: a) initiatives lacked
direction, b) the GMFRS personnel needed more support than just the toolkit with

evaluation activities, and c) GMFRS staff members found the knowledge gap, lack of resources and organisational guidance as barriers to carrying out evaluations.

These findings have highlighted the need for the following future actions, to be addressed in Cycle 3:

1. The importance of project management – embed a project planning model that can help GMFRS personnel clarify project goals, and which can underpin evaluation activities; and

2. An Evaluation Framework – to address the lack of organisational support for evaluation. Should include:

   - Policy – stating the role of evaluation in the organisation;

   - Staff – ensure staff are clear about their responsibilities and there is guidance and support available (in addition to the toolkit);

   - Standards – set quality standards for evaluation – what is GMFRS expecting the evaluations to look like;

   - Audit and Review – ensure evaluations are monitored and reviewed on a regular basis; and

   - Communicate – create channels to share the findings with wider organisation and stakeholders.

8. CYCLE 3 – DEVELOPING RECOMMENDATIONS FOR WIDER EVALUATION FRAMEWORK

In this chapter, details are given of the research activities carried out, and the findings made from the third Cycle of the action research study developing evaluation capacity to GMFRS. In Cycle 2, it was ascertained that, in addition to the toolkit, GMFRS staff needed additional assistance with evaluation activities; the evaluation trials showed that the participating staff members were not able to complete the different evaluation processes without guidance from the researcher, and the interviews also revealed that staff members regarded limited understanding of evaluation processes and methods, and insufficient organisational support as barriers to evaluation. The aim of the third Cycle was to address these findings by reviewing the new evaluation toolkit with users; and developing recommendations for a wider evaluation framework, investigating how the recommendations had been implemented, and what the organisation learnt from the process. The chapter is divided into five sections. The first describes the research methods used in the Cycle, the second summarises the findings made. The third examines the findings in the light of relevant literature, and in the fourth, the researcher reflects on the achievements of the Cycle. The final section contains a conclusion from this Cycle. The Cycle varies slightly from the previous two; the first two Cycles concentrated on examining the research problems and identifying possible approaches to address them, whereas this Cycle focuses on the change aspect of action research, and provides evidence of how the GMFRS evaluation practices were transformed during this project.

8.1 Description of Research Activities

The research activities of the third Cycle are a culmination of all the recommendations made in the previous Cycles. Cycles 1 and 2 emphasised the importance of embedding a project planning model that could help GMFRS personnel clarify project goals and underpin evaluation activities, and the need for a wider evaluation framework to address the lack of organisational support for evaluation. The Cycle started with a document analysis of meeting notes taken during a meeting with senior GMFRS personnel about the recommendations for a wider evaluation framework. This was followed by interviews with toolkit users and a focus group with various GMFRS staff members to review the final version of the toolkit. Figure 17 summarises the different stages of Cycle 3.
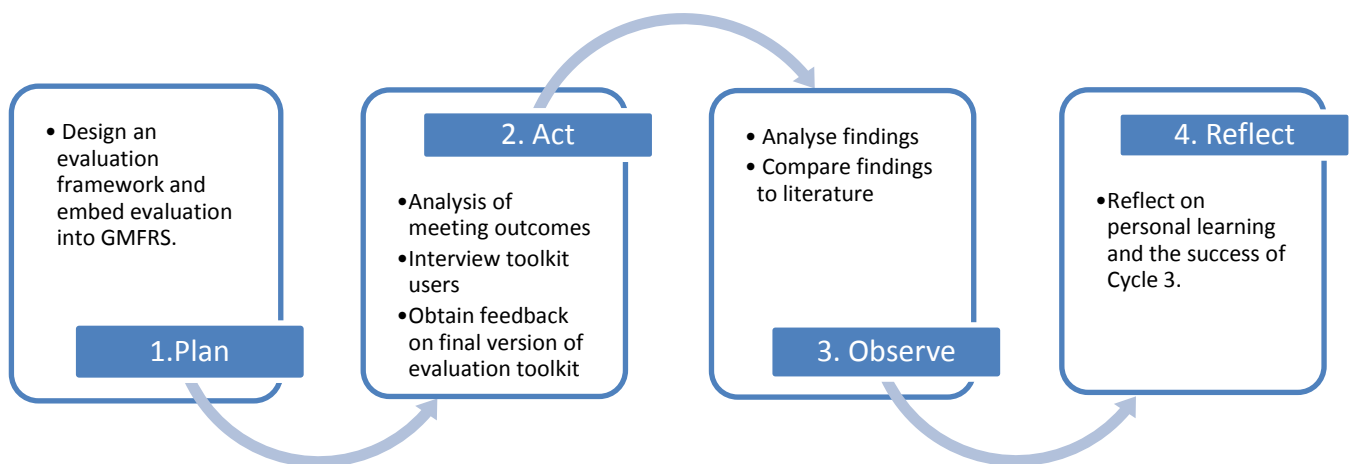

Figure 17: Cycle 3

Analysis of Meeting Outcomes

A recommendations paper addressing all the findings made during the previous cycles was drafted between November 2009 and January 2010 to help to gain

organisational approval and support for additional evaluation processes. The document was aimed at the Brigade Management Team (BMT), the body responsible for overseeing the running of the fire service, and contained details of the research processes, findings and a list of recommendations for a wider evaluation framework that would ensure that evaluation would become an embedded organisational process. Organisational approval and commitment to the recommendations was needed in order for the action research project to accomplish its aim: "the ultimate aim of the study is to develop a theoretical model for the use of evaluation capacity building (ECB) in a large public sector emergency service". The BMT meeting took place on the 19th February 2010. The researcher could not attend the meeting due to GMFRS regulations, but was represented by a senior staff member. Afterwards, a meeting about the BMT outcomes and the organisational commitment to the recommendations was arranged with the person. The meeting took place on the 24th of March 2010 in the staff member's office during normal office hours. The meeting was attended by the researcher and a GMFRS uniformed officer, meeting notes were taken by the researcher.

After an approval had been gained for the recommendations, the organisation demonstrated its commitment to evaluation by choosing to incorporate them into a new initiative management / problem solving methodology[10], SARA. SARA (Scanning, Analysis, Response and Assessment) is a commonly used problem-solving method associated with Problem-Oriented Policing (POP). POP is an approach to policing in which activities are subject to a thorough examination in order to discover new and more effective ways of dealing with them (Goldstein 2001) [11]. Several meetings were organised during March and May 2010 to discuss how the different evaluation processes could be

---

[10] New methodology to GMFRS
[11] See Appendix G for information about SARA.

linked to the new project management methodology. The meetings were attended by the researcher, and two non-uniformed staff members in charge of the SARA developments, and the electronic design of the documents. The meetings took place during normal office hours in the GMFRS Headquarters' meeting rooms; the researcher used an open note technique to record the meetings.

Interviews

The aim of the evaluation trials, organised in Cycle 2, were to test the usability of the new evaluation materials, but as described in Chapter 7, the participants tested the evaluation materials and processes set out in the toolkit in Cycle 2. These interviews were carried out with the staff members who had participated in the trials, to seek feedback on the final version of the toolkit, and to investigate what the participants thought of evaluation after they had had some experience of it. The participants were selected because of their roles in the trial process; they were all in managerial roles and had provided the most assistance with the evaluations. Five staff members were invited, and three of them agreed to participate. The interviewees were all uniformed members of staff, and the interviews were carried out in March 2010, in the participants' offices, during normal office hours, and they were conducted by the researcher. At the beginning of the interviews, the participants were given a copy of the interview questions, and they were also verbally repeated by the interviewer during the interview. Notes were taken by the researcher in front of the participants.

<u>Focus Group</u>

A focus group was organised to create awareness of the evaluation toolkit and to seek feedback on the final design. The session was divided into two parts; in the first part the researcher introduced the final version of the toolkit, including how to use it and how it had been designed; and in the second part the attendees were asked to answer a set of questions about layout and content of the product (Appendix H). The focus group was held on the 8th April 2010, during normal working hours in the GMFRS Headquarters' meeting room. The attendees were selected with two GMFRS senior managers, and initially invitations were sent to fifteen staff members. The invitees included both uniformed and non-uniformed GMFRS officers of various ranks, responsible for the future evaluation activities of the organisation. The focus group was attended by eleven staff members. A GMFRS officer hosted the discussion, and the researcher acted as a non-participant observer.

Notes were taken in two ways: a GMFRS administrator was asked to take minutes of the meeting, and record the decisions made. These notes were circulated to all attendees after the meeting. The researcher also observed the situation, and kept personal notes of the participants' views and behaviour.

8.2 Findings

Document Analysis

The following recommendations were made to the BMT to allow the implementation of the evaluation toolkit and the embedding of the culture of evaluation into the organisation[12]. The recommendations were part of a report that contained details of the research project and its key findings.

1. Policy – GMFRS requires an organisational statement demonstrating on-going support for evaluation and highlighting the importance of evaluation activities. The policy should address, and influence all the activities related to evaluation: the role of evaluation in the initiative management system; the roles and responsibilities of staff members intending to evaluate projects; the use of resources; and the way and when, work is designed, carried out, and monitored;

2. Organisational Arrangements - To make the policy effective, GMFRS staff members need to be involved in and committed to evaluation. GMFRS needs to ensure staff members are clear about their responsibilities and there is guidance and support available for those undertaking evaluations. Key issues to consider:

   • Competence: training and advisory support for all staff members.

   • Control: management process that allows continuous monitoring of evaluation activities and the quality of evaluations.

---

[12] Recommendation structure adapted from HSG65 model (Health and Safety Executive 1998)

- Communication: communications channels for data sharing.

- Co-operation: clear process allowing data sharing and cooperation with other Fire and Rescue Services and external stakeholders.

- Planning and Setting Standards – Planning is a key part of evaluation, and has to be incorporated into the initiative planning process. Evaluation planning involves setting objectives, choosing performance indicators and developing evaluation materials.

3. Standards should identify who does what, when and with what result. Standards must be measurable, achievable and realistic, and state what GMFRS is expecting the evaluations to look like. In order for the staff to produce reliable and meaningful reports, evaluations have to follow the same standards;

4. Audit and Review – Evaluations need to be monitored and reviewed on a regular basis, to ensure quality standards are met and that the reports show an objective picture of the initiatives. Monitoring should involve regular inspections and checks to ensure evaluations follow all the guidelines set in the policy and the quality standards; and

5. Learn and Communicate – Utilisation of evaluation findings is an essential part of evaluation. If the findings are not used to improve the initiative or to make decision about the future of the project, the evaluation has been a waste of resources. It is also equally important to communicate the findings, negative or positive, to the wider organisation and all stakeholders.

The initial feedback after the meeting (based on emails received from two senior officers and reproduced below) proved that the organisation had accepted the recommendations:

19 February 2010

Heidi

The tool kit was very well received today. You are to be congratulated on a very well put together document. Your observations on what is needed next were acknowledged and accepted and add weight to the developments in train regarding our BIKM programme and the efficacy of using POP[13] as a vehicle to embed the toolkit. Thank you for your efforts and attention to detail and the patience that you have needed at times during the project's development.
Kind regards, (Name withheld)
19 February 2010

Colleagues

Just to add to (name withheld) comments - the toolkit was really good and provides a really helpful platform to develop initiatives that have real impact. The early design work also looks very impressive indeed

(Name withheld)

A meeting was organised with the officer who represented the researcher at the BMT meeting to discuss the outcomes further. In the meeting, the researcher was looking for additional evidence of change and organisational willingness to commit to evaluation. During the meeting, the officer confirmed that the toolkit would be incorporated into the

---

[13] Problem-Oriented Policing

SARA methodology, and the organisation had made plans to employ an Evaluation and Researcher Officer to offer further guidance for GMFRS staff. Employing an evaluator and embedding evaluation processes into a project management methodology addressed all of the recommendations made for the organisation. These steps ensured evaluation would become an embedded process, monitored and utilised to its full potential. The officer also showed two documents to demonstrate the new direction the organisation was aiming to take. In Figure 18 it is shown how the document detailing the goals of the organisation had been redesigned to include aims, objectives and outcomes[14].

---

[14] Previously only included aims

Our purpose is to protect and improve the quality of life of the people in Greater Manchester

Prevention	Protection	Response	Public Value	People	Principles

What we aim to do

Engage with Greater Manchester's communities to inform and educate people in how to reduce the risk of fires and other emergencies and do all we can to prevent crime and disorder

Influence and regulate the built environment to protect people, property and the environment from harm

Plan and prepare for emergencies that may happen and make a high quality, effective and resilient response to them

Manage risk through using resources flexibly, efficiently and effectively, continuously improving our use of public money in ways the public value

Work with people with the right skills and attitude to deliver high quality, value for money services in a positive environment for everyone

Operate in accordance with the law and our values, and ensure that safety, sustainability, partnership and diversity run through all our activities

How we aim to achieve this (the means)

Prevention
- Targeted marketing, advice and education to change behaviour
- Helping communities to help themselves
- Investigating the causes of fires and helping to detect crimes
- Understanding the nature of our diverse communities
- Learning from incidents
- Monitoring, evaluating and improving our effectiveness

Public Value
- Research
- Risk management, planning and performance management
- Utilising our brand, expertise and resources beyond our statutory duties
- Maximising the use of our resources
- Using technology to improve services and efficiency
- Supporting economic growth
- Involving our communities in designing, evaluating and delivering our services
- Promoting the nature, purpose and value to the public of our organisation
- Being influential locally, regionally, nationally and internationally to benefit Greater Manchester

People
- Best practice approaches to recruiting people
- An integrated approach to workforce planning, succession planning and career development
- Effective learning and development programmes and facilities
- Maintaining competence and continued professional development
- Our people taking appropriate responsibility for their development and maintenance of competence
- Providing consistent assessment and appraisal
- Recognising and rewarding good performance and dealing fairly and consistently with poor performance
- Informing, consulting and involving our people in issues that affect them

Principles
- Effective internal and external communication
- An internal management system audit framework
- Setting clear standards and accountabilities
- In accordance with our values, holding each other to account
- Seeking to work in partnership to achieve better outcomes than could be achieved alone
- Providing safe persons and a healthy and safe working environment
- Providing our services in a sustainable way
- Providing services tailored to meet the needs of our diverse communities
- Being an employer of choice for all our communities

Performance measures tracking progress towards our service delivery and service development goals

The outcomes of our aims

Prevention
- Fewer emergency calls
- Fewer deaths and injuries from fires and other emergencies
- Less crime and disorder
- Less property damage and economic loss
- Less damage to the environment
- High public satisfaction with our services

Response
- High state of preparedness for emergencies
- Effective emergency cover and high quality response to emergencies
- Resilience during peak activity levels and business disruption
- People rescued from harm
- Fewer deaths and injuries from fires and other emergencies
- Less property damage and economic loss
- Less damage to the environment
- More business resilience and economic growth
- Our heritage preserved
- High public satisfaction with our services

Public Value
- Better designed, high value for money services
- Improved quality of life outcomes for communities
- More economic growth
- Fire stations at the heart of communities
- Volunteers adding further value to our service offer
- High public satisfaction with our services

People
- High quality, well trained workforce
- High performing workforce
- Effective workforce cover
- Committed and enthusiastic staff working in a positive environment
- A well informed workforce
- Low levels of ill discipline, absenteeism and staff turn over
- High levels of staff and volunteer satisfaction

Principles
- Better outcomes for our communities using the same resources but through partnerships
- A healthy workforce
- Low rates of accidents and injuries to our people
- Equitable service delivery by a workforce that is representative of the communities it serves
- Less impact on the environment from our services
- High levels of staff and volunteer satisfaction
- High public satisfaction with our services

Figure 18: Improved Organisational Goals

The officer had also planned a new activity to promote GMFRS to the community, and developed an initiative plan to include descriptions of achievable objectives and

outcomes, something that, as discussed in Cycle 1, had not previously been done (Figure
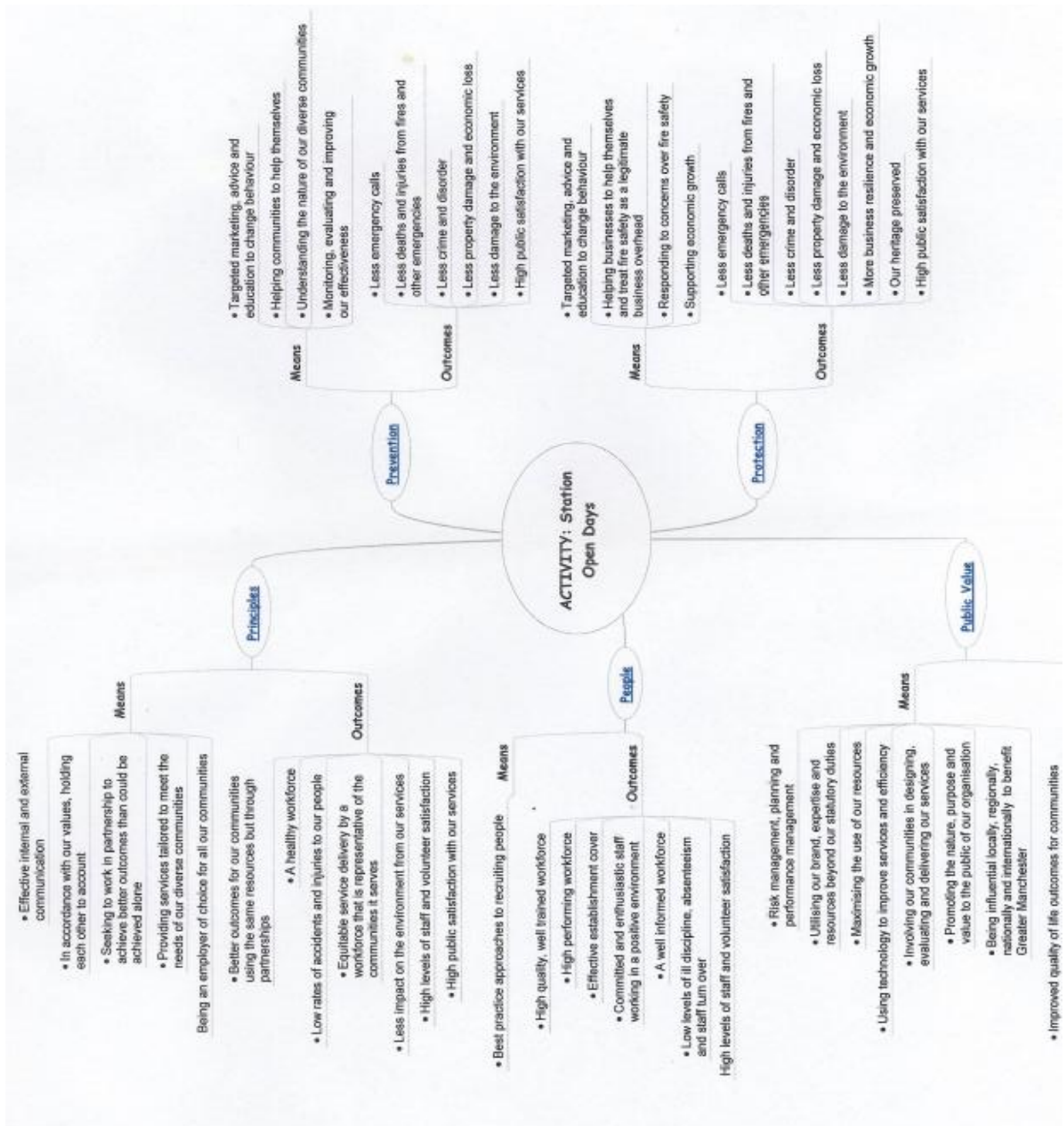
19).

Figure 19: Aims and Objectives of Station Open Days Initiative

After an organisational approval had been gained for the recommendations, several meetings were organised to discuss how the different evaluation processes could be linked to the SARA methodology. SARA, an essential part of the POP process (explained in more detail in section 8.1), is a problem solving process that will aid the planning and monitoring of GMFRS activities and partnerships. The methodology allows GMFRS to determine the most appropriate response to emerging problems, as well as assists in the planning of the effective actions, providing a complete project management methodology, that supports the use of evaluation, for GMFRS. The outcomes of these meeting are demonstrated in Figures 20 and 21. Figure 20 details how the different stages of evaluation correspond to the SARA methodology, and how the evaluation toolkit can be used to guide users through the different stages of evaluation. Figure 21 demonstrates what the new online database, that provides GMFSR staff members access to project management and evaluation resources, will look like.
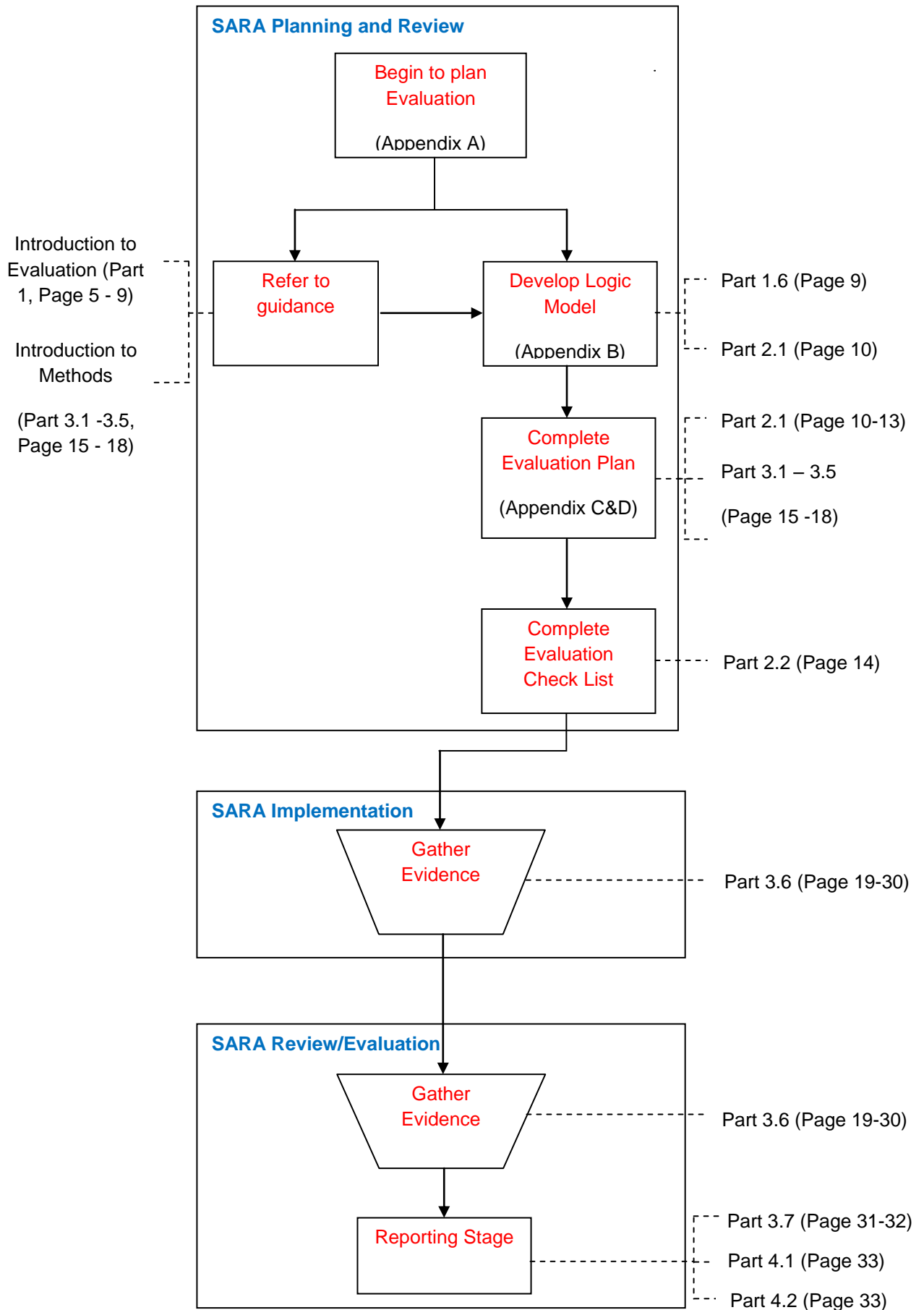
Figure 20: Evaluation Process Diagram

One of the key findings from Cycles 1 and 2 was that the evaluations were not utilised in anyway. GMFRS did not provide staff with any channels to communicate and share findings, and staff members thought that hindered the value of the existing evaluations. Hence, one of the key recommendations in the BMT recommendations paper was to create channels to share evaluation findings with the wider organisation and stakeholder. In the recommendations paper, it was argued that utilisation of evaluation finding was an essential part of evaluation. If the findings are not used to improve the initiative or to make decisions about the future of the project, the evaluation has been a waste of resources (Patton 1997). Figure 21 sets out the plan for an online database that guides users through the SARA process, including evaluation. The database also works as a repository for project documents and evaluation findings to increase the possibility of GMFRS staff using them for learning purposes.
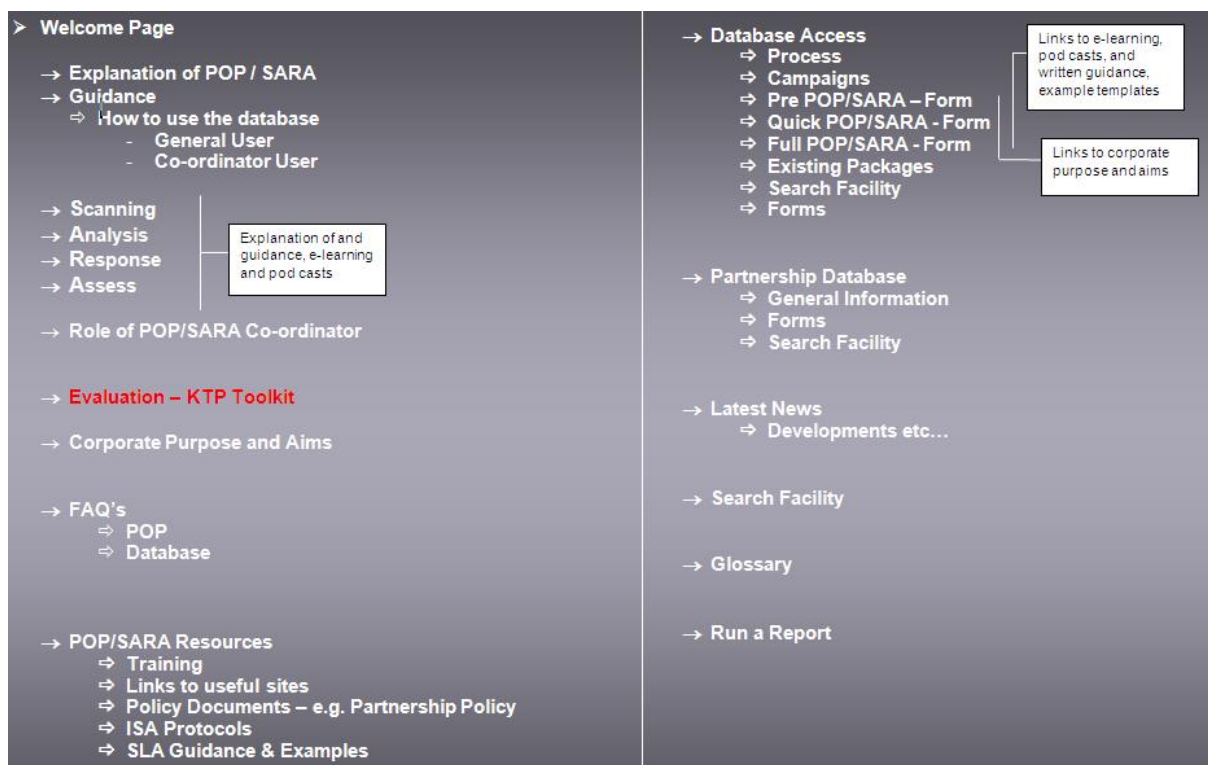
Figure 21: Snapshot of Database – Initial Outline Bullets

<u>Interviews</u>

The interview questions were divided into two parts: questions about evaluation processes and questions about the design of the toolkit (Appendix I). The questions asked in the first part of the interview were very similar to those asked in the Cycle 2 interviews. The only difference was that the participants were also asked if they had noticed a difference in their attitudes to evaluation after taking part in the evaluation trials. First, the participants were asked what they thought was the purpose of evaluation in this organisation, and if their views had changed since taking part in the evaluation. All of them reported very positive views about evaluation, and like the Cycle 2 interviewees, were able to describe the key aims of evaluation:

> "You can get feedback on how well the project has run. Need to take a lot of things into consideration, good or bad, resources, personnel – internal and external. If you want to improve a project, evaluation is vital – will give you all the answers."
> *Participant 1*

> "Firstly you can justify what you are doing and based on how it's performing. If it's achieving what it claims. Gives evidence. GMFRS has a history of doing things just for the sake of doing things."
> *Participant 2*

> "To see if it is working, GMFRS putting time in, best value, what do we get back – is scheme worth the time? Are aims and objectives being met? Allows changes to the project."
> *Participant 3*

Two interviewees thought their views had not changed since taking part in the evaluation trials, and one explained views had become more concentrated:

"Views have come more concentrated. First I felt it was a good thing to do, but since first time I met you I realise I had to underpin what I'm doing, to realise what to look at. Not only a subjective idea of what I'm doing."
Participant 2

When asked if they thought evaluation had any disadvantages, two of the participants, mentioned, similarly to Cycle 2 interviewees, resources:

"Time – takes a lot of time. Info gathering is very time consuming, and the more people you need for the evaluation the more difficult it gets. Difficult to get people engaged – can't spend time on chasing people up – and then decide if their views are important."
Participant 1

"Apart from time and filling in forms, not really. It has to be done. It's better to evaluate on a regular basis, if you do that all the time, allows you to check all the aims – makes the project dynamic."
Participant 3

One of the participants had a very different view of the disadvantages of evaluation, not witnessed in prior research activities:

"Organisational level, because there isn't an evaluation culture. People talk about evaluating. It seems like it's used as a manager's threat. By having a structured evaluation embraced by everyone - cannot use it as a threat."
Participant 2

This view was interesting, as all the other GMFRS staff members who had taken part in the action research project felt that the organisation did not provide support for evaluation activities, and it was not a systematic requirement or a process. The participants were also asked to feedback on the clarity of the toolkit and if they believed it was fit for purpose. Participant 1 found the toolkit very valuable, and mentioned that:

"It's simple and sets out all the steps. Good starting point. Very valuable and every initiative should use it. From beginner's point of view this toolkit is vital for beginners, very easy to read – don't want a document that needs a law degree to be able to understand/read it. Gives an indication of different areas you need to use like resources. Easy to follow – no need for training and suits everyone, normally FRS just asks people to do things without them having any knowledge of the issues."
Participant 1

The participant did not want GMFRS to organise any training to use the evaluation toolkit, but thought it would be useful to have a person to contact if any questions arouse. Another one of the participants found the toolkit very clear and easy to follow, but was hoping for more detail about data analysis:

"Easy to follow. Looks good. Not sure about the detail when it comes to the depth of analysis. I'm coming from psychology background - not sure about the qualitative side of things."
Participant 2

And the final participant had a more critical view of the product and thought it contained too much detail:

"When you read it, different people would pick different things. Had to read it twice – maybe too much detail, tick box things very useful. It wasn't clear, key words – used a lot of words that were only defined at the back of the document. Big document- can't just flick through."
Participant 3

The person also thought assistance might be needed when using the toolkit for the first time:

"Back up call would be useful. Maybe training - but needs to be evaluated to ensure its cost effective - the way GMFRS has trained people before doesn't work. Senior manager gets trained, and everyone assumes he/she will then train his/her staff. Anyway If I was to evaluate, someone would evaluate it again, because there is no trust."
Participant 3

<u>Focus Group</u>

The participants were categorised according to their rank to examine if their position in the organisation had any impact on the number of contributions they made. A very senior uniformed officer participated in the focus group, and the researcher had discussed with GMFRS staff members close to the project, the possibility of that having an impact on the lower rank participants. It was feared that they would not feel comfortable voicing their opinions in the company of senior management. However, as Table 8 shows, the rank of the officers did not affect the number of contributions made. On average participants made 3 comments, and there was little variation between the ranks. Another factor that could have affected the number of contributions was the fact that some of the participants had an opportunity to view the document prior to the focus group. The conversation break down reveals that participants who had previous contact with the toolkit were slightly more likely to contribute to the conversation. When investigating the number of contributions made, an interesting pattern was witnessed. Nearly every other comment was made by A, the senior officer. Participant A did not host the focus group, but an investigation of the conversation patterns revealed that A felt obliged to address all comments made by lower rank staff members made. This could have been due to the seniority of A's role, and the culture of the organisation.

Table 8: Focus Group Conversation Patterns

Rank	Participant	Number of comments	Had seen the toolkit prior to the focus group
Senior Management	A	8	No
	B	2	Yes
Non-uniformed	C	0	No
	D	4	Yes
Middle Rank	E	7	No
	F	2	No
	G	3	Yes
	H	2	No
	I	0	No
Fire Fighter	J	4	Yes
	K	2	Yes

After the toolkit introduction, the participants were asked if the toolkit was easy to use, clear and covered all aspects of evaluation in enough detail, and if they would like to change anything about the document. All of the participants were supportive of the toolkit, and two participants already knew someone in the organisation who was planning to use it in their work (not related to community initiatives). No one had anything to add to the toolkit, or wanted to make changes. Three of the eleven participant commented how they found it flexible to use:

"Think so – idea was to create something that could fit all and everything and hopefully people will be more encouraged to do evaluations. Sets out process very clearly."
Participant F

"Toolkit sufficient for every evaluation."
Participant B

"Yes it is flexible (toolkit), but needs a phone number on it to provide guidance. Essential to have guidance with the actual evaluation."
Participant J

Following these comments there was a discussion of how the toolkit had addressed the problem of initiatives having no aims and objectives. Three participants agreed that the toolkit forces people to questions why they carry out their activities, and helps them to put them into logical order:

"When Heidi started no SMART objectives. Long term aims are common, but short term aims not specific to the project. Toolkit forces you to question why I'm doing this."
"Our scheme hadn't had properly defined objectives. Problem: didn't understand evaluation. Toolkit enables to put thinking into logical order."
Participant E

"Using boxing initiative as an example – this piece of work has been at numerous meetings, but couldn't define why boxing club was good. Moss Side [boxing initiative] hasn't achieved original aims and objectives – but because evaluations weren't clear. To have this toolkit will provide clarity to those. Will it work? It's essential it to be part of the commissioning of a project. If people don't make it / force people to use it will not be used. People not comfortable to declare what has been achieved – need to explain what they are going to deliver."
Participant A

"p.9 [of the evaluation toolkit] logic model is the most essential part of the document. As far as inputs go we evaluate, mostly we are completely blind to outputs – very weak area. Often we have an eye for what outcomes are, but we don't have evaluation questions. We don't know what outputs are because we don't have clearly defined evaluation questions. If we haven't done logic model we will end somewhere unknown (sic)."
Participant A

The final question was "In addition to the toolkit, should the organisation provide any further assistance with evaluation?" In a very brief conversation four participants made a contribution, and they mainly showed concern about the interim management of evaluation activities until the SARA methodology becomes a fully embedded process:

"In 18 months' time we will have everything set for SARA and evaluation (including guidance), but we need more support with the toolkit /evaluation in the interim process. Could develop library of examples from PIM process."
Participant D

"GMFRS had limited knowledge of documents and evaluation – [name withheld] and [name withheld] will act as contacts."
Participant E

"Communications is a big gap in this organisation. How do Fire Fighter level get all the necessary information?"
Participant J

"But one thing – BMT changes – responsible for corporate planning and performance (in the future) –top end corporate planning – and it occurs to me that I will fold the toolkit into that process. The SARA coordinator should act as the evaluation officer."
Participant A

8.3 Review

The aim of this Cycle was to develop a wider evaluation framework for GMFRS, and to ensure the final version of the evaluation toolkit matched the user's requirements. It was established in Chapter 3, that ECB approaches should take the following issues into consideration: 1) individual's ability to conduct evaluations, which in this study was accomplished by developing an evaluation toolkit to guide users through the different

stages of evaluation, 2) the organisational capacity to use evaluations, which was tackled by embedding evaluation into the SARA methodology, and 3) the various stages of building the necessary processes that accommodate and support both individual and organisational capacity to evaluate, addressed by the recommendations for a wider evaluation framework, including an evaluation policy, employment of an evaluator, and audit and review process to ensure the quality of evaluations. During the literature review, it was noted that there was a need to step away from the thinking that ECB is only an activity that teaches people to evaluate (Williams 2001 in McDonald et al. 2003, Huffman et al. 2008, Hay 2010). The argument advanced by Williams (2001 in McDonald et al. 2003) that purely focusing on building evaluation skills can lead to a situation where an organisation may be capable of producing evaluations but unable to use them, or even worse, produce evaluations that are treated as irrelevant, is very accurate in the GMFRS context. The organisation had previously tried to employ an individual ECB approach i.e. encouraged staff to use evaluation materials and hope it would become an embedded process; however as this study has shown, without the organisational support, and a wider evaluation framework, the aim of developing evaluation capacity into the organisation would not have been achieved. A review of ECB literature identified that many of the current ECB efforts concentrated on developing individual skills and expertise (Stevenson et al.2002, Lennie 2005, Monroe et al. 2005, Cohen 2006, Forss et al. 2006, Taut 2007, Adams and Dickinson 2010).

Some authors had tried to move away from the individualistic and expert centred approaches by recommending ways to implement evaluation capacity into the organisations, and describing structures that are essential in developing sustainable evaluations systems (Minnett 1999, Barnette and Wallis 2003, McDonald et al. 2003,

Wandersman et al. 2003, Arnold 2006, Miller et al. 2006, Huffman et al. 2008), Taylor-Powell and Boyd 2008 and Adams and Dickinson 2010). This action research study has confirmed that focusing purely on individual's ability to produce evaluations does not build an organisational capacity to evaluate. All the research findings have showed that, in addition to individual capacity, an organisation has to build processes to support evaluation activities. To compare the GMFRS ECB project to other organisational ECB approaches, such as the ones carried out by Wandersman et al. (2003), Arnold (2006), and Taylor-Powell and Boyd (2008), it can be noted that the organisational recommendations have many similarities: provision of evaluation assistance and resources, leadership, structures, policy and procedures to support the use of evaluation. However, the results of this study also indicate that evaluation cannot exist in isolation alongside other organisational processes, it has to be embedded into them; something that previous ECB efforts have failed to acknowledge. For example, in the case of GMFRS, into a project management system to ensure the organisation produces quality evaluations, as well as utilises the findings to improve projects, and informs stakeholders of the project's achievements.

8.4 Reflections

<u>Reflections</u>

The third Cycle involved three key activities: making recommendation for wider evaluation framework, embedding evaluation into the organisational process, and gaining feedback on the final version of the toolkit. Taking into consideration the challenges faced in Cycles 1 and 2, and the lack of interest in evaluation and the action research study, the findings of the BMT meeting and the focus group were immensely positive. It was a relief

and a thrill to witness senior staff members discussing, and acknowledging, the lack of aims and objectives the community initiatives had, and making plans for embedding evaluation activities into the organisation. For the first time in the action research project's history, the organisation had provided evidence of its commitment to evaluation. The researcher had spent 18 months trying to voice concerns about the way the initiatives were run, and the lack of organisational processes to support evaluation, and thought developing ECB for GMFRS was impossible. But finally the organisation showed that the recommendations were taken seriously, and they started to realise the benefits of evaluation.

All of the one-to-one meetings worked very well, but the focus group could have been changed to an alternative research method. During the focus group it occurred to the researcher that the lower rank GMFRS staff members might have never been asked to provide feedback on processes / tasks they would have to carry out, because some of the staff members seemed unsure of how to behave. The lower rank officers were cautious of voicing their opinions, and the more senior members felt like they had to take control of the conversation, even though the idea of the focus group was to treat everyone equally. The focus group lasted an hour, and the atmosphere was slightly uncomfortable; this could have also been due to the presence of senior staff, or the general culture of the organisation. Unfortunately there was not enough time to run two separate focus groups, or use research methods that focus more on the individual, such as interviews. In terms of methods suitable for groups the Delphi method could have been used, even though it would have limited the free flow of the conversation. In hindsight, the focus group was probably the best method, but the participants should have been divided into smaller and rank appropriate groups.

With regards to the literature, all of the ECB approaches that had discussed the importance of involving the whole organisation in the ECB effort gave very brief and vague descriptions of the organisation's role in the ECB process. This could be due to the fact that all of them had been developed by external ECB developers, which meant the companies were left to deal with the organisational aspect of ECB. The most inspirational studies were McDonald et al. (2003) and Taylor-Powell and Boyd (2008) because they had made contributions to both, to the ECB definitions debate, and included in-depth analyses of individual and organisational approaches to ECB. They had both considered various other ECB approaches before developing their own, however, as with the other authors, they provided very little evidence for their arguments. The majority of the ECB studies have been carried out after 2000, and as it is a relatively new field, many issues remain unresearched, including the questions raised in Chapter 3. But also, as noted by Sanders (2002) and Cousins et al. (2004), the field needs more empirical studies, as the majority of the ones found in the literature are only theoretical. Also a wider variety of contributions would enhance the development of the field. As noted previously, all, apart from one, studies have been written from the external evaluators' perspective – it would be interesting to include alternative perspectives in the debate.

Learning

- Prior to starting an action research project, identify how to gain support of senior management;
- When trying to change an organisational process, provide the organisation with tangible evidence of the benefits of the product or the intended change process.

The saying "you can't build a reputation on what you are going to do" applies to many of the frustrations experienced during the project; and

- Take the culture of the organisation into consideration prior to choosing research methods. Contemplate carefully the use of group methods in hierarchical organisations.

8.5 Conclusion

This chapter contains details of all the research activities carried out during Cycle 3, and the findings made. The Cycle consisted of three different types of research activities: a document analysis of meeting notes taken during meetings with senior GMFRS personnel about the recommendations for wider evaluation framework, interviews with toolkit users and a focus group to review the final version of the toolkit. The activities of this Cycle were carried out for three reasons: to gain feedback on the final version of the toolkit, to improve initiative management and recording keeping system, and to investigate how the recommendations for a wider evaluation framework were received and implemented. This Cycle has provided evidence of successfully incorporating evaluation processes into the organisation's initiative management process.

9. DISCUSSION

In this chapter the results of the action research study developing evaluation capacity to GMFRS are discussed with reference to the findings of the literature review. The chapter is divided into five sections. In the first the GMFRS ECB project is summarised, in the second, the findings are discussed in the light of the literature, and an analysis of new themes that have emerged from the research is provided. The third recommends topics for future research, and in the fourth, the researcher reflects on the achievements of the action research study. The final section concludes the chapter by summarising its content.

9.1 Summary of GMFRS Project

The ultimate aim of this action research study was to develop a theoretical model for the use of evaluation capacity building (ECB) in a large public sector emergency service. The study was divided into three Cycles, which all dealt with a different aspect of ECB. In the first Cycle, the concentration of the research activities was on understanding the organisation's evaluation needs. The findings of this Cycle showed that the community initiatives, and their recording / management structure lacked direction, which had resulted to poorly defined and evaluated initiatives. These findings were essential for the future Cycles, because the lack of understanding about the project's goals has important implications for evaluation - clear and explicit project objectives are a precondition for assessing the effectiveness of a service (Phillips et al. 1994). This prompted the researcher to start investigating how to improve the initiative management system which would accommodate systematic evaluation practices. A study of the existing evaluation tool also

revealed that it was not detailed enough for use by personnel with no previous evaluation experience. In the second Cycle, a new evaluation toolkit was developed and tested, and interviews were conducted about barriers to internal evaluation practice at GMFRS. The findings of the second Cycle reiterated the findings of the first Cycle: initiatives lacked direction, and that GMFRS personnel needed more support than just the toolkit with evaluation activities. In addition, GMFRS staff members found the lack of knowledge about evaluation activities, lack of resources and organisational guidance as barriers to carrying out evaluations. The aims of the research activities carried out in the final Cycle were to develop a wider evaluation framework, and embed evaluation into the organisation. The overall themes that centred every Cycle were: poorly defined projects, lack of understanding/guidance/resources to evaluate, and a need for organisational commitment to evaluation and all necessary resources. The following section of this chapter will examine the key findings in the light of ECB literature.

9.2 Addressing Knowledge Gaps

In this section, the key findings of the action research study on developing evaluation capacity for GMFRS are discussed and compared to other ECB approaches found in the literature. The current ECB approaches can be divided into individual and organisational approaches. The individual approaches develop staff members' skills by training them to use research methods for evaluation, and logic models to assist in evaluation planning (Stevenson et al. 2002, Monroe et al. 2005, Lennie 2005, Cohen 2006, Forss et al. 2006, Taut 2007, Adams and Dickinson 2010). In this study, the need for individual skills is addressed by developing an evaluation toolkit that covered all the different stages of evaluation. The research activities of Cycles 1 and 2 indicated that

GMFRS staff had a poor understanding of evaluation activities; limited skills and knowledge about evaluation at all levels in organisations is common, hence it is important to demystify evaluation (Duignan 2003). The first section of the evaluation toolkit provided an overview of evaluation, and answered questions such as, why, when and what to evaluate, to ensure the organisation had a shared understanding of evaluation and its use. The second section addressed the need to clarify project goals prior to an evaluation. It guided the user through completing a logic model, evaluation plan and a timetable for the evaluation. The aim of the second part was to introduce a structured approach to planning an evaluation, something that previous GMFRS evaluation activities had been lacking. The third section was about gathering evidence and making sense of findings, it explained to users how to choose the right type of methods for the assessment of their initiatives, and described how to use the research methods, and analyse the data. The final section guided users through composing the right kind of evaluation report to summarise the findings of the evaluation, and explained how to utilise and share the evaluation findings with the rest of the organisation and all necessary stakeholders. All the different stages were supported by examples of completed plans and evaluations to help users through the process. The individual ECB approaches developed by Stevenson et al. (2002), Miller et al. (2006), Naccarella et al. (2007), and Taylor-Powell and Boyd (2008) also included the provision of external assistance to complement the evaluation materials and training. GMFRS did not have resources to provide external assistance or organise training events, and there was a belief that staff might forget the skills / knowledge if not used immediately after the training, and by the time they came to evaluate, they would need to complete the training again. The interviews carried out in Cycle 3 indicated that the toolkit was so self-explanatory, that additional training was not required, and for those staff members who wished for one-to-one guidance on evaluation processes or the use of the toolkit contact

details of two staff members, that were familiar with evaluation activities and the toolkit approach, were provided.

Prior to the action research study, GMFRS had already tried, without any success, the individual approach to ECB. Individual approach refers to the training of individual members of the organisation to use evaluation materials without the organisation providing any additional support in the evaluation process. The research activities of this study also demonstrated that the toolkit / evaluation training / individual approach was not enough to build the organisation's capacity to evaluate. Cycle 1 showed that the initiative's goals needed some clarification in order to carry out an effective evaluation, and Cycle 2 revealed that staff needed more assistance with evaluations, as well as support and resources from the organisation. GMFRS had no system in place to help staff in the design, conduct and utilisation of evaluations. There was also no incentive to evaluate, as it had not been made mandatory, and staff members were unclear about how and when to carry out evaluations. This confirmed the arguments advanced by Huffman et al. (2008) and Hay (2010) that there is a need to step away from the thinking that capacity building is purely an activity that teaches people to evaluate. In the case of GMFRS, the organisation needed both a framework that supported the use of evaluation, as well as organisational commitment to the activity of evaluation. Williams (2001 in McDonald et al. 2003) summarised the situation perfectly, where GMFRS would have ended if the ECB approach had only included the individual approach, when stating that all the skills, knowledge, technical expertise and experience in the world will not help, if the programme, community, organisation, or environment cannot sustain and nurture those skills and abilities.

The work carried out by Minnett (1999), Barnette and Wallis (2003), McDonald et al. (2003), Arnold (2006), Miller et al. (2006), Huffman et al. (2008), Taylor-Powell and Boyd (2008), and Adams and Dickinson (2010) highlighted the need to engage the whole organisation in the ECB process, since if the organisation does not support and integrate evaluation it will not have an evaluation system, which essentially is the focus of ECB. The key argument that separates these authors from the authors of the individual approaches is that ECB has to address all the organisational aspects required to develop an evaluation system that allows individuals, with the support of the organisation, to carry out meaningful evaluations. Evaluation capability should provide enduring organisational benefits, including a sustainable resource for producing evaluations, as well as a system for encouraging and using evaluation (Duignan 2003, Preskill and Boyle 2008, and Taylor-Powell and Boyd 2008). The organisational approaches include the provision of leadership, evaluation assistance, resources (time, money, software), communication channels to utilise evaluation findings, and policy and procedures that set the "rules and regulations" of evaluation. These findings were confirmed by this action research study. The researcher developed a wider evaluation framework to support the use of evaluation, and ultimately build the organisational capacity to carry out, and utilise evaluations. The recommendations are summarised below:

- Policy – stating the role of evaluation in the organisation;
- Staff – ensure staff are clear about their responsibilities, and there is guidance and support available;
- Standards – set quality standards for evaluation – what are you expecting the evaluations to look like;

- Audit and Review – Ensure evaluations are monitored and reviewed on a regular basis; and

- Communicate – create channels to share the findings with wider organisation and stakeholders.

The recommendations addressed all of the findings made during the research, but most importantly ensured the organisation commits to evaluation, and provides clear guidance and management to the staff members wishing to carry out evaluations. The policy will state the role of evaluation in the organisation – when, how and by whom they will be carried out, as well as the quality standards each evaluation has to adhere to. The audit and review, as well as assistance and training needs, are met by the employment of evaluation officers who will ensure evaluation activities follow the correct procedures, and staff receive all the necessary guidance. These recommendations were very similar to the other organisational approaches, however as the focus of ECB is on developing processes and practices that make evaluation part of the everyday work of an organisation (Huffman et al. 2008:359), they were not enough to ensure the routine use of evaluation. In Chapters 3 and 6 it was demonstrated that evaluation looks at the achievements of a project, and how and why these have occurred. Clear, specific and measurable programme goals are part of good evaluation conditions (Phillips et al. 1994, Patton 1997). GMFRS had problems in communicating the goals of the initiatives in a way that they would have been helpful in evaluation. Project management deals with planning, coordinating and controlling projects (Lock 2007), and the different evaluation activities correspond to the stages of project life cycle, as shown in Table 9. The project planning stage should set the aims and objectives of the project, as well as an evaluation strategy because monitoring and evaluation are key aspects of project control (Cleland 1999, Longman and Mullins

2005, Meredith and Mantel 2006). Planning an evaluation is as important as planning any other aspect of a project (Cleland 1999). Therefore, a recommendation was made to incorporate the different stages of evaluation into the daily project management activities of the organisation. During the Cycle 2 interviews, GMFRS staff had also identified the lack of resources as one of the biggest barriers to evaluation, however when evaluations are planned at the project initiation stage, staff can identify their resource needs prior to the evaluation activates starting. Embedding evaluation into project planning and management also ensures all evaluation activities are timed and spread throughout the project cycle, and turn it into a routine activity.

Table 9: Phases of Evaluation Activities Corresponding to Project Cycle

	Project Planning Stage	Project Implementation Stage	Project Termination
Process Evaluation	- Decide to evaluate - Define purpose, timeline, resource requirements and budget for evaluation - Decide methods for data collection	- Develop / refine methods for data collection - Collect data - Analyse data - Write the report - Decide what to do with the project – change direction etc. - Communicate findings with the organisation and all the key stakeholders	
Outcome/Impact Evaluation	- Decide to evaluate - Define purpose, timeline, resource requirements and budget for evaluation - Decide methods for data collection	- Develop / refine methods for data collection - Collect data	- Analyse data - Write the report - Decide what to do with the project – carry on, change, etc. - Communicate findings with the organisation and all the key stakeholders

Boyle and Lemaire (1999), McDonald et al. (2003), and Preskill and Boyle (2008) have also discussed the need to embed evaluation capacity into the organisation in order to develop sustainable evaluation systems that are systematically utilised. According to Boyle and Lemaire (1999) and Preskill and Boyle (2008), the organisational support functions will automatically turn evaluation into systematic practice. In contrast, McDonald et al. (2003) argue for making evaluation compulsory, even though they acknowledge that "the most important lesson… is that performance management (and evaluative enterprise in general) cannot be forced on people. Attempting to impose it will likely lead to goal displacement, unreliable information and an increase in the risk that programme relevance will be diminished rather than augmented" (Davies, 1999:157 in McDonald et al. 2003). GMFRS did not have a culture of creating and using information effectively, and without embedding evaluation into the project management process, GMFRS would have ended up in a situation where the organisation may have been capable of producing evaluations but unable to use them, or even worse, produce evaluations that are treated as irrelevant (Williams 2001 in McDonald et al. 2003). In Chapter 3, after an examination of various meanings attributed to ECB, it was concluded that evaluation capacity consists of the individuals' ability to carry out evaluations as well as the organisation ability to provide the necessary functions required producing evaluations. Furthermore, building evaluation capacity was defined as building a system that allows an organisation to have the ability to perform and produce functional evaluations; meaningful and functional evaluations are well planned, and systematically produced and utilised with organisational support. For a learning organisation, a sufficient ECB approach may mean the development of organisational support systems. However, in the case of GMFRS, the organisation would have not had the capacity to evaluate if, in addition to developing all the functions, evaluation had not been made embedded, systematic and easily accessible.

The literature review also identified a gap in knowledge regarding the sustainability of the ECB models. Naccarella et al. (2007), Stevenson et al. (2002), Forss et al. (2006) , Huffman et al. 2006, Taut 2007, and Compton (2009) argued that the use of external assistance would allow an organisation to maintaining the ECB structures, but in the GMFRS case, the organisation did not have the funds for this kind of resource. However, the organisation's willingness to accept the recommendation made by this study, and incorporating evaluation into the SARA methodology, addressed this gap. Every project that is planned using SARA has to consider evaluation. The methodology allows evaluation to become an automatic consideration, and the ECB effort to become sustainable in the long run. With regards to barriers GMFRS staff found the lack organisational guidance, as well as lack of resources and knowledge about evaluation, hindered the evaluation activities of the organisation. The researcher found the lack of project planning and organisational commitment as barriers to ECB. Without linking project planning / management and evaluation together, the study would not have succeeded in the aim of embedding evaluation into the organisation.

This study has enhanced the field of ECB by providing new knowledge about how to, not only build, but to embed evaluation capacity into a major public sector emergency service. To summarise the arguments made in this chapter, Figure 22 was developed to represent the ECB model developed during this study.

Figure 22: ECB Model

Figure 22 details the individual and the organisational responsibilities in the ECB process. The smaller 'individual' circle is placed inside the bigger 'organisation' circle to represent the relationship between the evaluator and the organisation. ECB is a joint effort between the two circles, and consists of the individuals' ability to carry out evaluations as well as the organisation's ability to provide the necessary functions required to produce evaluations. As noted in Chapter 2, ECB consists of the individuals' ability to develop evaluation models, decide meaningful evaluation questions, collect useful data and make sense of the data together with the organisational ability to utilise the data in a way that it benefits the project, all necessary stakeholders and the organisation. The third circle titled as 'project management' represents the need to embed evaluation capacity into a project

management framework to: a) ensure evaluation activities are incorporated into the project's life cycle from the start; and b) make evaluation a sustainable activity. As discussed in the preceding paragraphs, GMFRS – a non-learning organisation - did not have a culture of creating and using information effectively, and without embedding evaluation into the project management process, GMFRS would have ended up in a situation where the organisation may have been capable of producing evaluations but unable to use them, or even worse, produce evaluations that are treated as irrelevant (Williams 2001 in McDonald et al. 2003).

9.3 Recommendations for Further Research

While engaging with the literature, and carrying out this action research study, ideas for further research emerged. The ideas could be divided into ones that further the field of ECB, and those that enhance the evaluation practices of GMFRS. Starting with the literature related suggestions, it was found that the literature lacks empirically tested approaches (also identified by, Sanders 2002 and Cousins et al. 2004, Preskill and Boyle 2008), and very little is known about the organisational context of the ECB studies. More data about the context would allow comparisons between different models, and the examination of possible commonalities between approaches that had been developed for similar types of organisations. ECB is very context specific, hence a new model is created for each study, but an increase in the number of context studies would allow the identification and mapping of structures and approaches that suit certain types of organisations. This would be particularly beneficial at the practitioner level.

The field of ECB is relatively new, but the researcher believes that rather than purely developing new models, the field should also deepen its knowledge of the existing ones. Deeper knowledge also relates to the other idea for further research – evaluation of current approaches. It was observed during the literature review, that none of the current ECB approaches have been evaluated, which, again, raises questions about their impact and sustainability. There is a clear gap in the literature about the benefits of ECB to organisations, and how the evaluation practices have been maintained in the long run. The researcher believes all the ECB models, including this action research study, should be evaluated few years[15] after the ECB effort to study the impact of both the ECB approach, and evaluation on the organisation, and an investigation of questions such as:

- do ECB approaches create systematic evaluation practice?;
- has the organisation witnessed an increase in the number and quality of evaluations after the ECB effort?;
- what has the organisation learnt from the ECB experience and the evaluations?; and
- has the ECB effort equipped an organisation to produce usable evaluation?

This would be very beneficial in the case of GMFRS as their expected outcomes for this research project were to develop an evaluation toolkit, and to embed evaluation and feedback into the organisations community initiative design process. To ensure the newly developed evaluation capacity is used to its full potential, the organisation needs to be proactive and evaluate the benefits of its evaluation capacity system.

---

[15] Depending on the organisation

The other ideas for further research relate to the organisation. During the literature review, it was identified that to strengthen evaluation practices beyond methods, collaborations with evaluation experts, and organisational arrangements, a deeper shift in the culture and mind-set is needed for evaluation practice to prosper (Milstein et al. 2002). This means that in order for GMFRS to fully benefit from the capacity to evaluate, and the actual evaluations, the organisation has to embed it into its culture which, ultimately, makes evaluation into a mainstreamed activity. Some approaches, to allow the GMFRS ECB to develop into an evaluation culture and ultimately to a mainstreamed evaluation practice, are identified by Porteous (1999) and Wandersman et al. (2003) and Cousins et al. (2004). Cousins et al. (2004), argue that it is achieved through sustained evaluative inquiry and, in particular, continued and routine use of evaluation findings and processes. When evaluation becomes integrated into the on-going activities within an organisation, it may become a learning system that fosters the development of shared values and understanding among organisation members (Cousins et al. 2004). Porteous (1999) and Wandersman et al. (2003), on the other hand argue for the use of Fetterman's empowerment evaluation approach. Empowerment evaluation is "an evaluation approach that aims to increase the probability of achieving program success by (1) providing program stakeholders with tools for assessing the planning, implementation, and self-evaluation of their program, and (2) mainstreaming evaluation as part of the planning and management of the program/organization" (Wandersman et al. 2005:28). Empowerment evaluation is achieved by linking evaluation into capacity building and social justice (Fetterman and Wandersman 2007). Also useful are Sanders' (2002:256) set of indicators to assess if evaluation has been mainstreamed within an organisation, something that GMFRS could use as goals for their efforts:

- Someone asking "how are you going to evaluate this?";

- Board members asking for evaluation findings for every program report;

- Evaluation appearing on the agenda of every staff meeting;

- Buyers asking for evaluation data from every sales agent;

- Clients being asked for evaluation that is then taken seriously by staff members;

- The CEO distributing a list of organizational values that includes continuous evaluation;

- Orientation training for new employees that includes their role in evaluating services, policies, and products; and

- Using evaluation advocacy as a selection criteria when hiring new staff.

These indicators would ensure evaluation becomes an automatic consideration at every level of the organisation. The individual and organisational processes developed during this research project will ensure the organisation has the capacity to evaluate and the right process in place to support the evaluation activities. But in the future evaluation has to be brought to the fore front of organisational thinking that the capacity is developed into a learning system.

9.4 Reflections on the Action Research Study

Reflections

The goal of the study was to make a contribution to both practice and theory: to equip GMFRS with the knowledge and processes to enable them to carry out effective and systematic evaluations, and while doing so develop new knowledge about ECB to enhance

theory. The study was carried out through an action research approach, because the methodology allowed the focus of the research to be on changing a process, and me to have the dual role of an employee and a researcher. The research resulted in developing a framework that improved the GMFRS' community initiatives evaluation and planning practices.

Being an active part of the GMFRS helped to understand the organisational culture, the views people held about evaluation, and to gain an understanding of their needs in terms of evaluation practices. The methodology was well suited to the situation, as the dual role of employee and researcher assisted in gaining an insight into the organisation that, in my opinion, an external researcher would not have been able to gain. GMFRS had already tried the external evaluator approach for developing capacity to evaluate, but without major success. Many of the ECB authors, such as Stevenson et al. (2002), Forss et al. (2006), Huffman et al. (2006), Naccarella et al. (2007), Taut (2007), and Compton (2009) also emphasise the importance of participatory approaches to ECB, and the researcher believes it was essential to be part of the organisation, and not be seen as an "outsider", as well as to involve the staff members in the research, because it helped the organisation to gain ownership of the project, and evaluation practices as whole.

The method selection was based on whatever data was available at the time. Reflecting on it now, in an ideal world it would have been good to have had more input in the selection, but then that would have defeated the aim of action research, as argued by Dick (1993) and Stringer (2007), action research represents the unknown situation under investigation; hence, it cannot be started with specific research questions, as the flow of the inquiry will determine them and the solutions to the research activities have to be

found from the context. The chosen methods did serve their purpose, and were representative of the researcher's abilities, resources available, and the research environment. It would have been useful to use more one-to-one interviews, but due to the participants, and the project's busy schedules, it was not possible. A lot of data was also disregarded from the thesis because it had been very poorly recorded, an indication of the demanding timescale, and hence not suitable for academic research.

Overall the project was very successful and yielded many unexpected outcomes in terms of organisational and personal learning. During the first two Cycles many challenges were encountered, and questions raised about evaluation and ECB, and it was sometimes impossible to see a successful end to the research project. All the reflections that have followed each Cycle demonstrate the new knowledge and skills the organisation and the researcher have acquired from the research. Without having to go through all the stages of developing ECB myself, I would not have learnt as much as I have from the project.

9.5 Conclusion

In this chapter the results of the action research study developing evaluation capacity to GMFRS were discussed with reference to the findings of the literature review. The first part of the chapter discussed the findings in the light of the literature, and provided an analysis of new themes that emerged from this action research study. The first section summarised the study, the second recommended topics for future research, and in the third, the researcher reflected on the overall achievements of the research project. This study has argued that the current ECB approaches are designed for external evaluators, not for organisations hoping to develop their own evaluation capacity, skills and expertise. The

current ECB approaches utilise a variety of methods to build organisational evaluation capacity and describe the functions needed for successful ECB. However, none of them have discussed how to embed, and sustain, the evaluation skills and processes to ensure they are systematically utilised. There was also a gap in the knowledge of how to embed evaluation capacity into an organisation that does not have the funds or capacity to accommodate the additional support functions/organisational processes associated with ECB. This study has provided a new empirically tested ECB approach to a non-learning public sector emergency service with limited resources, and detailed the development of evaluation capacity from the staff members' perspective.

10. CONCLUSION

The ultimate aim of the study was to develop a theoretical model for the use of evaluation capacity building in a large public sector emergency service. It was identified in the literature review that the ECB literature lacks theoretical models and/or guidance to assist emergency services and/or non-learning organisations with no previous experience of systematic evaluation practices to develop internal evaluation capacity. By developing ECB for GMFRS, this action research study has addressed that gap, and enhanced the field by providing knowledge of how to, not only build, but to embed evaluation capacity into a major public sector emergency service – an organisation with limited funds and resources, and which cannot be categorised as a learning organisation. The findings of the study enhance the field of ECB by providing a new project management focused ECB approach, and also offer other non-learning, and emergency services that do not have existing processes in place to create and use knowledge effectively, an empirically tested ECB approach. This new approach does not only assist in developing individual and organisational processes to evaluate, but in addition recommends how to embed evaluation into the daily activities of an organisation. The research has also produced data that detailed barriers to developing ECB in a non-learning organisation, and the organisational support processes required to nurture individuals' evaluation skills and abilities.

The aim of the study was achieved by the following objectives:

1. to assess the process of:

   a) developing an evaluation framework and toolkit, and

b) supporting the embedding of an evaluation culture in the GMFRS against the original aims and objectives of the project, and comparing the findings to a relevant theoretical framework;

2. to investigate the challenges and benefits of non-experienced personnel using the evaluation toolkit successfully; and

3. to assess the success of a project designed to develop a self-evaluation framework and toolkit for use by non-specialists in the GMFRS.

The first two objectives were addressed in Cycles 1 and 2. In Cycle 1, the initiatives register and the existing evaluation materials were analysed and the findings of the research activities revealed that the community initiatives lacked clearly defined aims, objectives and outcomes, which had important implications for evaluation, as clear and explicit project objectives are a precondition for assessing the effectiveness of a service (Phillips et al. 1994). It was also established that the organisation's existing evaluation processes were not detailed enough for use by personnel with no previous evaluation experience. These findings highlighted the need to improve organisational understanding of evaluation, embed a project planning model to help to clarify the goals of the initiatives, and to develop new evaluation materials. These actions formed the basis of Cycle 2, in which an analysis of other evaluation toolkits was carried out, a focus group held, the use of the new evaluation materials observed, and staff interviewed about barriers to internal evaluation practices. The key findings of Cycle 2 were that the GMFRS personnel needed more support than just the toolkit with evaluation activities, and that a knowledge gap, lack of resources and organisational guidance were identified as barriers to carrying out

evaluations. These findings prompted the researcher to develop recommendations for a wider evaluation framework that would address the importance of a project management model that would underpin evaluation activities.

In the final Cycle the last objective was addressed in order to fulfil the aim of the research. Cycle 3 consisted of three different types of research activities: a document analysis of meeting notes taken during meetings with senior GMFRS personnel about the recommendations for wider evaluation framework, interviews with toolkit users and a focus group to review the final version of the toolkit. The activities of this Cycle were carried out to investigate how the recommendations for a wider evaluation framework had been implemented, to assess the organisational learning from the process, and to provide evidence of successfully incorporating evaluation processes into the organisation's initiative management processes.

To summarise and conclude the thesis, this study has detailed all the different stages of developing individual and organisational capacity to conduct and utilise evaluations, and advanced an argument that for evaluation to become an embedded and systematic activity, it has to be supported by a project management methodology that underpins evaluation processes. The study has developed a new evaluation system for a non-learning emergency service in which evaluation has previously failed, which will improve the business intelligence needs of a major emergency service, and allows them to make more informed decisions about the expenditure of scarce funds.

Appendix A – A Section of the Communication Plan

Ser No.	1. Stakeholder / Group	2. Message – i.e. what stakeholder needs to know	3. Method of Communication	4. Date	5. Frequency of Communication	6. Person responsible for communication	7. Information / feedback the stakeholder can give us
1	GMFRS	Introduction	Intranet / Journal Article	Dec	Once	Heidi , Comms dept	
2	Borough Commanders	Introduction	GRIP	25/9/08	Once	Heidi, ▮	General comments about the project
3	Supervisors	Progress	Meetings	Last Fri of month	weekly ▮ monthly	Heidi	Progress Review
4	4.1 SMT	Intro and Outcomes of Project stage 1	SMT meeting	Nov		Heidi, ▮	Comments about the project, suggestions for future activities
	4.2 Key Stakeholders (3 BC's)	Outcomes of Project stage 1	Email, Meeting	Nov		Heidi, Comms dept	
	4.3 GMFRS	Outcomes of Project stage 1	Intranet blog, and newsflash email	Dec			

Appendix B – A List of Research Methods for GMFRS Focus Group

Research Method	Hard Outcome (Method suitable for measuring quantifiable outcomes)	Soft Outcome (Method suitable for Measuring non-tangible / qualitative outcomes)	Process (Method suitable for process evaluation)	Impact (Method suitable for impact evaluation)	Cost Effectiveness (Method suitable for measuring cost effectiveness)
Focus Group		✔	✔		
Nominal Group Technique		✔	✔		
Interviews		✔	✔	✔	
Document Analysis / Service Utilisation	✔	✔	✔	✔	✔
Questionnaires and Surveys	✔	✔	✔	✔	
Diaries		✔		✔	
Observations		✔	✔	✔	
Goal Setting / Portfolio		✔		✔	
Test and Scales		✔		✔	
Cost - Allocation Cost - Benefit Cost - Effectiveness	✔				✔

Appendix C – Interview Questions – Barriers to Evaluation

1. What do you think is the value/purpose of evaluation?

2. Do you think the GMFRS should evaluate its own community initiatives?

3. Did you know there was an evaluation toolkit available to use at GMFRS?

4. Are you involved in any evaluation at the moment? If yes, in what way?

5. Do you think that the evaluations that are currently undertaken by the GMFRS are of value
   a. to GMFRS?
   b. to the initiatives?

6. In your opinion does GMFRS provide sufficient resources to assist the staff with planning and carrying out evaluations?

7. Do you think the GMFRS personnel have a clear understanding of the purposes and processes of evaluation?

8. Do you think GMFRS personnel need further training to carry out evaluations?

9. Do you think there are any constraints to evaluation in GMFRS?

10. How would you like to see the evaluation processes develop in the future?

Appendix D – Evaluation Toolkit

Greater Manchester Fire and Rescue Service

Evaluation Toolkit

GREATER MANCHESTER
FIRE AND RESCUE SERVICE

PREVENTING PROTECTING RESPONDING   www.manchesterfire.gov.uk

Contents

List of Tables

Acknowledgements

How to Use the Toolkit

This guide is designed to assist in the evaluation of Greater Manchester Fire and Rescue Service (GMFRS) community safety initiatives. It has been divided in the following sections, to guide the users through the different stages of evaluation:

Part 1 Understanding Evaluation

A brief introduction to evaluation. Designed to answer questions such as, why, when and what to evaluate.

Part 2 Planning Your Evaluation

Assists with evaluation planning and the completion of the evaluation plan. Also gives brief examples of completed evaluation plans.

Part 3 Gathering Evaluation Evidence and Making Sense of Findings

Guidance on evaluation methods and how to use them.

Part 4 Reporting and Sharing Evaluation Findings

Explains how to write an evaluation report, and how to utilise and share the evaluation findings.

Next Steps

If you are new to evaluation read Part 1 and work your way systematically through the rest of the document.

If you are familiar with evaluation activities, you may wish to use Appendix A as a reminder. If you want to refresh any of the points mentioned in the evaluation summary, please refer to the corresponding part of the document.

This guide offers a brief introduction to evaluation. If you require more guidance on the different evaluation processes or the use of any of the evaluation tools, please contact:

Terri Byrne Telephone: 0161 736 5866 Email: byrnet@manchesterfire.gov.uk
Ben Levy Telephone: 0161 736 5866 Email: levyb@manchesterfire.gov.uk

Part 1 – Understanding Evaluation

1.1 What is Evaluation?

The term evaluation can be used in a variety of ways. The following definition represents the approach recommended for assessing community safety initiatives delivered by GMFRS.

"Evaluation is concerned with judging merit against some yardsticks. It involves the collection, analysis and interpretation of data bearing on the achievement of an organisation's goals and programme objectives."
(Phillips 1994)

Evaluation refers to the activities that are used to explore projects and their achievements against the original aims and objectives. Evaluation includes four key stages: planning, gathering evidence, reporting and sharing findings.

Evaluation can provide evidence about how effective projects have been, and the reasons why. It is an essential component of effective and efficient project delivery, and enables the continuous improvement of community safety activities.

1.2 Why Evaluate?

Evaluation is a necessary part of all activities, as it helps to identify projects' effectiveness and reveal areas for improvement. Evaluation can help you to make informed decisions about why an activity is effective or ineffective and if it is an appropriate use of resources. It can also help to highlight future improvement needs, and determine if the project is worth delivering again, and if it is transferable to other areas.

1.3 Who Evaluates?

The person responsible for managing the project should ensure that evaluation is embedded in the project work plan.

An external evaluator could be considered when the initiative is resource intensive, highly visible, or an internal evaluation has indicated that a more in depth and/or objective assessment is necessary to gain further insight.

1.4 When to Evaluate?

Data collection is an important part of evaluation and has to be done during the project in a systematic way. Depending on the type of evaluation, the collection of data can take place at the beginning and end of the project, half way through, or throughout the project.

It is important to schedule evaluation activities into the project plan at the start of the project, to ensure that procedures are in place for the data collection at the appropriate time. If evaluation is only considered at the end of the project, it may not be feasible to collect meaningful data, in order to produce valid evaluation reports. It is also important to bear in mind that the collection of data might take place quite a while before and after the project to make sure any detected changes are real.

Table 1 indicates the approximate sequencing of the evaluation activities which take place during the different stages of the project life cycle:

Part One

Table 1: Phases of Evaluation Activities Corresponding to Project Cycle

	Project Planning Stage	Project Implementation Stage	Project Termination
Outcome / Impact Evaluation	• Decide to evaluate • Define purpose, timeline, resource requirements and budget for evaluation • Choose methods for data collection	• Develop / refine methods for data collection • Collect data	• Analyse data • Write the report • Decide what to do with the project - carry on, change etc • Communicate findings with the organisation and all the key stakeholders
Process Evaluation	• Decide to evaluate • Define purpose, timeline, resource requirements and budget for evaluation • Choose methods for data collection	• Develop / refine methods for data collection • Collect data • Analyse data • Write the report • Decide what to do with the project - change direction etc • Communicate findings with the organisation and all the key stakeholders	

1.5 Types of Evaluation

There are many different types of evaluations, however the three main ones are: Process, Outcome and Impact. In evaluation literature, process evaluation is also known as formative evaluation and impact as summative. All of the different types of evaluations can be used singularly or in combination.

Process

Process evaluation helps to establish if the project is running according to the initial plan. An effective project may not yield the desired results if it is not delivered properly. Process evaluation is normally carried out half way through the project to examine the inputs, outputs and short term outcomes, and it can inform you of how to change the delivery methods of the project in order to achieve the planned outcomes. The concentration is on:

- how initiatives / services are delivered
- what actually occurs while providing services
- strengths and weaknesses of delivery methods

Outcome

Outcome evaluation helps to identify what occurred as a result of your project. It determines whether short, medium and / or long term outcomes have been achieved. Outcome evaluations tend to concentrate on:

- the end results of the project and if the project met its overall goal(s)
- what happened as a result of the project

Impact

Impact evaluation goes a little further than outcome. It doesn't only measure the outcomes, but also looks at why and how the project has worked, and how much of the outcomes were caused by other events. Impact evaluation focuses on the final long term, intended and unintended results, and aims to identify what are the wider changes that have taken place as a result of the initiative. Impact evaluations concentrate on:

- why the particular project brought about the change / impact
- additional benefits and unintended outcomes

If the aim of the evaluation is to improve the programme, choose process evaluation, and if the intention is to prove it works – choose outcome / impact evaluation.

Table 2 lists the benefits of these types of evaluations.

Table 2: Benefits of Formative and Summative Evaluation Questions[1]

Process Evaluation - Improve	Impact Evaluation - Prove
Provides information that helps you improve your programme. Generates periodic reports. Information can be shared quickly.	Generates information that can be used to demonstrate the results of your programme to funders and your community.
Focuses most on programme activities, outputs, and short term outcomes for the purpose of monitoring progress and making mid-course corrections when needed.	Focuses most on programmes medium term outcomes and impact. Although data may be collected throughout the programme, the purpose is to determine the value and worth based on results.
Helpful in bringing suggestions for improvement to the attention of staff.	Helpful in describing the quality and effectiveness of your programme by documenting its impact on participants and the community.

[1] W.K. Kellogg Foundation "Logic Model Development Guide"

204

1.6 What to Evaluate?

Table 3 gives an indication of the issues that could be evaluated as part of process (see input/output/short term outcomes) or outcome evaluation (see outcome/impact). Impact evaluation is normally underpinned by a theoretical model, and in addition to output and outcome evaluations would also include an in depth study of the wider context and external influencing factors.

Table 3: Logic Model [2]

Inputs	Outputs		Outcomes - Impact		
	Activities	Participation	Short Term	Medium Term	Long Term
What we invest	What we do	Who we reach	What the short term results are	What the medium term results are	What the ultimate impact(s) are
Staff Volunteers Time Money Research Base Materials Equipment Technology Partners	Conduct workshops, meetings Deliver services Develop products, curriculum, resources Train Provide counseling Assess Facilitate Partner Work with media	Participants Clients Agencies Decision-makers Customers	Learning Awareness Knowledge Attitudes Skills Opinions Aspirations Motivations	Action Behaviour Practice Decision-making Policies Social action	Conditions Social Economic Civic Environmental

[2] University of Wisconsin Program Development and Evaluation 2002

Part 2 – Planning Your Evaluation

2.1 How to Plan an Evaluation?

1. Develop a Project Logic Model

To conduct an effective evaluation, the aims and objectives of the initiative must be clearly defined. If your project doesn't have clearly articulated aims and objectives, the project logic model (Appendix B) can be used to help you to clarify your thinking about what the project is designed to achieve, and which aspects of the project could be evaluated. Once you are clear about the outcomes the initiative is designed to achieve you can then start to think about the questions that you would need to ask to assess if these outcomes have been attained.

Example of a Completed Logic Model – Moss Side Boxing Club

Project Aim: To engage with local youth to prevent them from attacking firefighters and setting fires					
Inputs	**Outputs – Activities**	**Outputs – Participation**	**Outcomes Short Term**	**Outcomes Medium Term**	**Outcomes Long Term**
Building. Equipment. Time.	Teach boxing to 5 different groups, based on their skills and fitness levels. Educate young people about discipline, respect and courage to change.	Local Youths.	Community engagement. Young People - awareness of Fire Service activities, and positive attitude towards fire fighters. Young people - awareness of good behaviour.	Young people - increased self esteem, opinions of FRS changed, and behaviour improved. FRS - better understanding of community needs.	Less attacks on fire fighters. Less fire crime. Safer communities.
Assumption: Boxing has a certain kudos, it increases self esteem, and it can increase the respect that an individual gets from their peer group. If these young people have an influence in a gang or group situation and they have an affinity to the fire service then, we are in a win-win situation where we can help to form citizens of the future whilst having ambassadors for the Fire Service on the streets of Moss Side.				**External Factors:** Gangs GMFRS support Partner agencies Trainers availability	

Part Two

2. Develop an evaluation plan and select the audience

Select the people who are going to contribute to your evaluation. These could be FRS officers, other agencies or delivery partners, members of community and/or individuals taking part in the initiatives.

Develop an evaluation plan which should clarify what aspects of the project to evaluate, the questions to ask and the indicators of success. In Appendix C you will find a proforma to complete after taking account of the following:

How to choose the evaluation focus area?

Evaluation focus area refers to the part of the project you aim to evaluate. The focus area can represent the inputs, outputs or outcomes you are aiming to investigate in your evaluation. Depending on the type of evaluation and your resources, you can choose to evaluate all aspects of the project or only some of them. Decide on the focus area based on what you want to know about the project, and what resources you have available for the evaluation.

How to define an evaluation question?

Evaluation questions are the major questions related to each focus area - the questions you want to answer in the evaluation. Evaluation questions need to be relevant to the focus areas, and need to be kept as simple as possible.

How to select the indicators?

Indicators are observable and measurable milestones and/or targets that indicate to you whether the project or the participants are making any progress towards the outcomes.

For example if one of the outcomes of the project was to educate the participants about discipline, respect, and courage, the following table demonstrate the indictors that could be used to identify any progress made towards these goals.

Outcome	Indicators
Discipline	▪ turn up on time ▪ attend regularly
Courage	▪ don't give up ▪ courage to change and control aggressive behaviour
Respect	▪ respect others ▪ respect equipment ▪ listen coaches

How to choose the methods?

The definition of the evaluation questions and indicators are important precursors to the selection of data collection methods. The data-collection methods should be determined based on how appropriate they are for answering your key evaluation questions and for achieving the ultimate purpose of the evaluation. The credibility and usefulness of the evaluation results can be strengthened by mixing evaluation methods where appropriate.

Tie method selection to available resources, and what is appropriate for the target population. Bear in mind that data collection and analysis can be very time consuming; Part 3 can help you to determine suitable methods for your evaluation.

Example of a Completed Evaluation Plan – Moss Side Boxing Club

1. Evaluation Focus Area	2. Evaluation Question	3. Indicators	4. Method
Input Resources	What resources are used for project delivery?	Cost of building, equipment, travel and time	Data analysis and interviews
Output 1 Engage with local youth	Is the project reaching its target population?	The attendees' postcodes	Postcode mapping and analysis
Output 2 Mentoring - teaching the participants about the qualities of firefighters: Respect, Discipline and Courage.	What have the children learnt at the boxing club? What do the children get out of the boxing club?	Attitudes and beliefs of the participants (in relation to the questions) Discipline: turn up on time, attend regularly Courage: Don't give up, courage to change and control aggressive behaviour Respect: respect others, respect gym equipment, listen to coaches.	Interviews with trainers and participants Interviews with parents, schools and partner agencies Analysis of attendance records

3. Decide who collects the data, how often, when and where

To ensure the evaluation causes minimum disturbance to the project delivery, it is essential to plan the activities and everyone's responsibilities in advance.

When designing your evaluation it is important to consider the resources requirements. Plan for time and costs – evaluation is resource intensive, around 10% of the total project costs (including staff time) should be budgeted for evaluation. Appendix D can help you to design a timetable for the evaluation.

4. Design and Test Materials

Choose the methods you consider most suitable for your data collection from the list in Part 3. After you have designed your materials, it is advisable to test them on a target audience. Ideally, you should test them on the same kinds of people you will include in the study. However this is not always possible, so invite a few people to have a look at the tools to ensure they are easy to understand and give you meaningful answers.

Example of a Competed Evaluation Timetable – Moss Side Boxing Club

Moss Side Boxing Club																
Year 2009																
Task	**July**				**August**				**September**				**October**			
Training – Person A	X															
Interviews with participants and parents – Person B			X	X	X	X										
Interviews with trainers – Person A		X														
Interviews with partner agencies – Person B					X	X										
Postcode Analysis – Person A							X	X								
Document Analysis – Person B									X	X						
Data Analysis – Person A							X	X	X	X	X	X				
Report Writing – Person A													X	X		

2.2 Evaluation Check List

Before starting any evaluation process, make sure the following issues have been considered.

Item	Yes	No	Comments
Has the evaluation plan been completed?			
Have the right resources (time, money and staff) been allocated to the evaluation?			
Does the evaluation process have a clear timetable?			
Have all the stakeholders, including the participants, been informed of the evaluation?			
Have you addressed any emerging ethical issues, discussed in Part 3?			

Part 3 – Gathering Evaluation Evidence And Making Sense Of Findings

3.1 Introduction to Methods

There are a range of methods that can be used to gather the data on which an evaluation can be based. Each of the different methods have their own strengths and weaknesses which influences the ways in which they can be applied. Evaluations should make use of primary and secondary research, as well as a mix of different methods to increase validity and reliability. The methods that have been chosen as most appropriate for use in the evaluation of GMFRS projects are:

- Focus Groups
- Nominal Group Technique
- Interviews
- Record/Document Analysis
- Questionnaires

See Section 3.6 for further details on these methods, including their advantages, disadvantages and cost implications, as well guidance on how to use them.

3.2 How to Select the Methods

Table 4 indicates the suitability of each of the methods for measuring hard and soft outcomes, and their applicability to process, outcome and impact evaluations. The aim of the table is not to restrict you to these techniques, but to recommend some popular and reliable ones that are used for evaluation. If you wish to experiment with other methods, ensure they are reliable ways of testing and measuring the project achievements.

Before deciding on methods always consider resource implications, the projects delivery approach and target audience. Also, if the project has delivery partners, or if other agencies have evaluated or are considering to evaluate the project, think about sharing data with them and utilising it in your evaluation.

Table 4: Research Methods for Different Types of Evaluations

Method	Hard Outcomes	Soft Outcomes	Process	Outcome/Impact
Focus Group		✓	✓	✓
Nominal Group Technique			✓	✓
Interviews		✓	✓	✓
Record/Document Analysis	✓	✓	✓	✓
Questionnaires	✓	✓	✓	✓
Cost Allocation, Cost Benefit, Cost Effectiveness	✓		✓	✓

Part Three

When selecting the methods, consider how credible the evaluation will be as a result of the methods that have been chosen. When deciding between various methods and instruments, ask the following questions:

Are the methods suitable for the population being studied and the issue being assessed? Think whether the target group has any special needs. For example, some youth groups might have difficulties reading, therefore face-to-face interviews might work better than questionnaires. When dealing with children and young people keep the tools simple and easy to follow.

Is the data collection tool valid? In other words, does it measure what it claims to measure? For example if you aim to measure participants attitude to fire safety, ensure the questions are relevant to the topic, and focus on the indicators you are trying to assess.

Is the data collection tool reliable? In other words, will it provide the same answers even if it is administered at different times or by different people? For example if two evaluators undertook an interview with the same individuals, the responses should be very similar.

3.3 How to Increase Validity and Reliability of Evaluation Findings?

Triangulation - a good way to increase the credibility of your evaluation findings is to use more than one method to collect your data. Evaluation designs should incorporate both qualitative and quantitative data collection methods whenever possible.

Sampling - is concerned with the number of participants taking part in the research. In most cases the number of people taking part in GMFRS initiatives is fairly low, therefore it would be advisable to use the whole group as a sample for the evaluation. If this is not possible, make sure you include a variety of people in the sample. Try to select individuals who are representative of the participants and don't just

pick people from one group, or people who are likely to give you the answer you want.

Control Group - a group of people that is matched as closely as possible with participants, but who are not participating in the project, can be used to validate the changes that occur in the participant's group as a consequence of the initiative. However it is not always possible or realistic to include a control group in the evaluation process.

3.4 Limitations to Evaluation

Always try to ensure your data collection tools are properly developed, your sample group is representative of the reality and you have interpreted the findings objectively. However bear in mind that all evaluations have their limitations. In the real world no one can guarantee that an evaluation is 100% valid and reliable due to various issues. Some common factors that contribute to the limitations are experience, skills, resources (time and money) and location.

It is important to reflect on the limitations of the evaluation in the final report. Talking about the limitations doesn't mean the report is any less valid, but it helps the reader to understand the context in which the project took place, and allows you to reflect on the quality of the data collection tools, the use of resources and the effectiveness of the evaluation process.

3.5 Ethical Considerations

Participants in evaluation have moral and legal rights, and it is important that when you are conducting an evaluation you do not violate these rights. You should try to ensure that your enthusiasm for getting answers does not lead you to pay less attention to the rights of the participants. There are a range of issues related to research ethics that you should consider[3]:

- **Voluntary Participation** – Taking part in the evaluation should never be compulsory. Beware that some people are more likely to volunteer than others and reliance on volunteers may introduce bias into your findings.

- **Informed Consent** - It is important to ensure everyone taking part in the evaluation understands the purpose of the study, how they were selected, and what will happen to the evaluation report, so that they can make informed judgments about whether to take part.

- **No Harm to Participants** - The evaluation process or the end report should never cause any harm to the participants, including being embarrassed. Questions which may upset or harm the participants should not be asked.

- **CRB / Vetting** - Appropriate vetting procedures including a check of criminal records may need to be carried out on any person involved in the interview of children. Parental consent must be obtained to interview children under the age of 16, and it is good practice to gain consent for young people under 18. (For further guidance contact the CYP department)

- **Confidentiality** – all details of the participants should be kept private by the evaluator and not disclosed to a third party. If confidentiality is guaranteed, do not share participants details with anyone, and ensure the participants are not identifiable from any published materials. There are certain exceptions to confidentiality, when dealing with children and young people. The CYP department can provide further guidance on the matter, specifically with regard to the disclosure of personal issues by children.

- **Data Collection and Storage** - Evaluation may gather sensitive data regarding the participants. Ensure you comply with data protection guidelines available on BigRed and protect participant's anonymity.

[3] Adapted from "Ethical Issues" Social Research Method.net

3.6 How to Use the Methods

In Section 3.6 details of the following methods and their use is provided:

Method A - Focus Groups (p20)

A semi structured group discussion to examine stakeholder's attitudes and opinions to services, concepts and ideas

Method B - Nominal Groups (p22)

A group technique that allows individuals to think about the issues themselves before the groups makes a joint decision about the topic. Used to identify areas for improvement and as a decision making / voting tool.

Method C - Interviews (p24)

One to one discussion with a stakeholder or participant to collect qualitative data about the participant's attitudes and opinions.

Method D - Record/Document Analysis (p26)

The study of existing documentation including: existing FRS and partner data, archives, statistics and any other existing research data.

Method E - Questionnaires (p27)

Questionnaires can be used to collect quantitative and qualitative data and to measure attitudes, values, personal experiences and behaviour.

Method F - Cost Analysis (p30)

Examines the cost of an initiative and the use of resources.
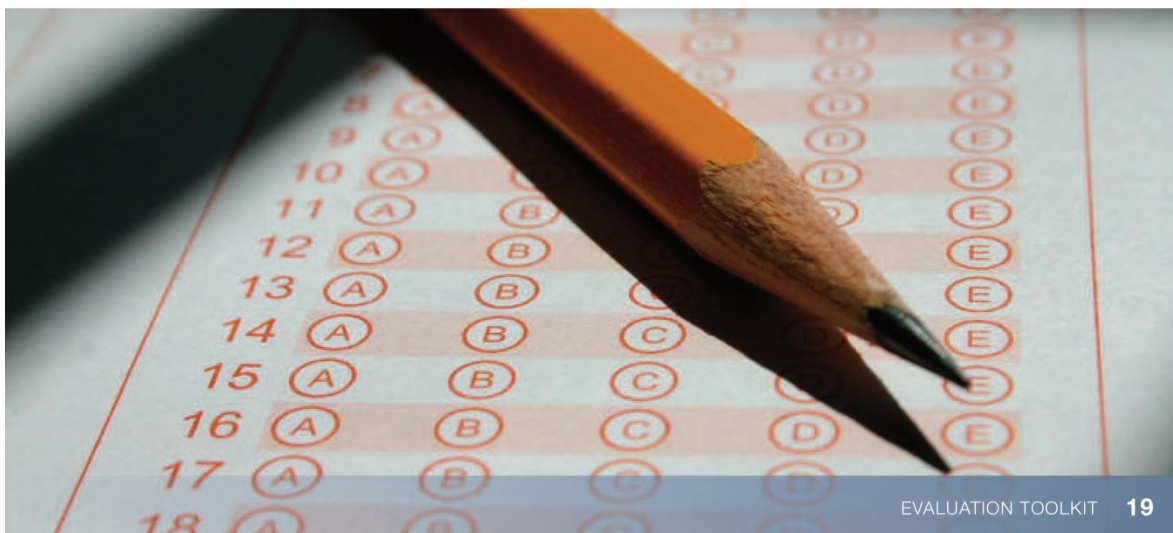
Method A – Focus Groups

Description

- A semi structured qualitative group discussion with 6 to 10 stakeholders

- Lead by a facilitator who follows an outline and manages group dynamics

- Focus groups are used to develop a better understanding of stakeholder's attitudes and opinions to services, concepts and ideas

- Questions are asked in an interactive group setting where participants are free to talk with other group members

Application

- Could be used as part of any evaluation process of any initiative

Advantages

- Provides in-depth information

- Focus groups have a high apparent validity - since the idea is easy to understand, the results are believable

- Low in cost; focus groups allow you to talk to several people at once and access results relatively quickly

- Limited time requirement on the participants

Disadvantages

- Participants can influence each other

- The evaluator has less control over a group than a one-on-one interview

- Time can be lost on discussion of issues irrelevant to the evaluation

- The data can be difficult to analyse because the discussion includes the participants reaction to the comments of other group members

Validity and Reliability

- Observer bias: the results obtained can be influenced by the evaluator, raising questions of validity

- Can be subjective due to the potential for facilitator bias

- Could be used in conjunction with other methods to increase validity and reliability

Resource Requirements and Cost

- Wages (session and analysis)

- Facilitator to set up and facilitate

- Observer to make notes

- Possible training for facilitators

- Audio equipment to capture discussions and play back for analysis

Analysis

- Principles of qualitative analysis

Additional Information

- Normally used in conjunction with observations, questionnaires and/or interviews

216

How to Run Focus Groups

Prepare a topic list and a question guide in advance of inviting around 6 to 10 people to participate in the focus group. It is advisable to record the discussion, but this is conditional on the informed consent of all the participants. Contemporaneous notes of the discussion should also be taken to supplement the recording.

A typical programme for a focus group is as follows[4]:

Introduction

- Explain what the topic is and why you are holding the focus group
- Ask the participants to introduce themselves and to prepare a name tag
- Explain the ground rules: confidentiality, everyone will have different views – respect that/don't criticise, everyone should have a say; everyone's views count

Sign in sheet

- To provide feedback if necessary

Content

- Work through the questions

Close

- Thank all the participants
- Agree to provide feedback

Tips for Running Focus Groups

- Develop your questions ahead of time

- Ask open ended questions

- Encourage free-flowing discussion and keep the session on track

- Start on an issue people are familiar with

- To encourage conversations, ask participants to think about an issue for a few minutes and write down their responses

- A summary document, should be produced for each session to circulate to participants

---

[4] Adapted from Tyne and Wear FRS Evaluation Guide

Method B – Nominal Group Technique

Description

- The nominal group technique is a decision making method for use among groups, who want to make their decision quickly

- Every member of the group gives their view of the topic. Duplicate views are eliminated, and the members proceed to rank the remaining ideas in order of preference. All scores are then totalled, revealing the most favoured items.

- It can identify strengths versus areas in need of development

Application

- Could be used as part of any evaluation process of any initiative

Advantages

- Effective decision-making

- Eliminates peer influence/ "group thinking"

- Produces large number of ideas in a short space of time

- Easy to analyse

Disadvantages

- Opinion may not converge in the voting process

- Ideas may be constrained

Validity and Reliability

- Observer dependency: the results obtained can be influenced by the evaluator, raising questions of validity.

- Subjective potential for facilitator bias.

- Could be used in conjunction with other tests to increase validity and reliability

Resource Requirements and Cost

- Wages (session and analysis)

- Facilitator, (Observer)

- Possible training

Analysis

- Principles of quantitative analysis

Additional Information

- To be used in conjunction with other tests (Triangulation)

How to Run Nominal Groups

In this technique, five to nine participants sit around a table, together with a leader. If there are more participants, they are divided into small groups. A single session, which deals with a single question, usually takes about 60-90 minutes. The basic steps are[5]:

1 Silent generation of ideas in writing – after making a welcoming statement, the leader reads aloud the question that the participants are to answer. Then each participant is given a worksheet (with the question printed at the top) and asked to take five minutes to write his or her ideas. Discussion is not permitted.

2 "Round-robin" feedback of ideas – the leader goes around the table and asks each member to contribute one of his or her ideas summarised in a few words and write them on a flip chart. These ideas are numbered and written so they are visible to all members. The process goes on until no further ideas are forthcoming. Discussion is not permitted during this stage.

3 Serial discussion of ideas – each of the ideas on the board is discussed in turn. The objective of this discussion is to obtain clarity and to air points of view, but not to resolve differences of opinion.

4 Preliminary vote – the participants are asked to select a specific number of "most important" items from the total list (usually five to nine). Then they are to rank these items on cards. The cards are collected and shuffled to maintain anonymity, and the votes are read out and recorded on a tally-chart that shows all the items and the rank numbers allocated to each.

5 Discussion of preliminary vote – a brief discussion of the voting pattern is now permitted. Members are told that the purpose of this discussion is additional clarification, and not to pressure others to change their votes.

6 Final vote – Step 4 is repeated.

Tips for Running a Nominal Group

- Develop your question ahead of time

- Encourage individuals to express their views at the appropriate time and keep the session on track

- Make sure you have flip chart sheets, the means of displaying them and the voting cards with you

- A summary document should be produced for each session to circulate to participants

[5] Kellogg Foundation Evaluation Handbook

219

Method C – Interviews

Description

- Interviews are used to collect qualitative data about the participant's attitudes and opinions

- Telephone or in-person one-on-one interviews

- Interviewer follows an outline but has flexibility

Application

- Could be used as part of any evaluation process of any initiative

- Interviews are particularly good to measure soft outcomes, especially if participants lack literacy skills

Advantages

- Eliminates peer influence

- Opportunity for interviewer to explore unexpected issues

- Can provide very detailed information

- Allows the use of probes, such as pictures, to measure attitudinal change

Disadvantages

- Can be very time consuming (interview and analysis)

- Potential for interviewer bias

- Interview skills required

- Can be difficult to analyse

Validity and Reliability

- Potential for interviewer and interviewee bias - the interviewer's own perceptions and question asking style can affect answers

- Poorly worded questions, ways questions are asked and misunderstandings can engender unreliable responses

- Validity and reliability issues could be limited by using structured questions and by undertaking interview skills training

Resource Requirements and Cost

- Wages (interview and analysis)

- Possible training

- Plenty of one-on-one time with the participants

- Plenty of time for the analysis, especially if the sample size is large

Analysis

- Principles of qualitative analysis

Additional Information

- Good for pre and post tests

How to Run Interviews

Interviews can give in-depth and detailed information, and can be used in all phases of an evaluation. The inside knowledge gained from interviews can provide an in-depth understanding of hard-to-measure concepts such as the impact of peer pressure on behaviour.

When interviewing participants/stakeholders, work through your question list and write down everything they say. Do not comment on their views, whether negative or positive, or empathise with their feelings. Your role is only to listen what they have to say and be respectful of their views. If at any point during the interview you are not clear about their comments, ask them to clarify. Do not be tempted to summarise their views based on what you think they might have said. Also allow the participants to see the questions and all the notes that you make. No secrecy is necessary, as it might harm the relationship with the interviewee.

If you wish to record an interview, first obtain permission from the interviewee. If there are indications that the presence of the tape recorder makes the Interviewees uncomfortable, consider taking handwritten notes instead.

Some Guidelines for Interviewing[6]:

Introduce Yourself

- Explain who you are and why you wish to interview the participants, how long the interview will take, and that taking part in the study is voluntary.

- Emphasise that the participants help is important and encourage them to speak openly. "We want to know what you really think!"

Confidentiality

- If anonymity is promised, it must be respected

- Offer them copies of your notes if they want them

The Interview

- Behave as neutrally as possible, even if the views that are expressed are not what you would wish. Pay special attention to your body language and tone of voice

Recording

- Keep writing as you talk, and if face to face keep as much eye contact as possible. When conducting phone interviews use encouraging/ listening noises or phrases

The End

- At the end thank them for their time, explain what happens next, how they can see the report and affirm the confidentiality

---

[6] Adapted from Kellogg Foundation Evaluation Handbook

Method D – Record/Document Analysis

Description

- Refers to the study of existing documentation, including FRS and partner data, archives, statistics and any other research data

- Aim is to collect information about participants or projects in a standardised manner

- Includes service utilisation - the use and utilisation of partner data

Application

- Could be used as part of any evaluation process of any initiative

Advantages

- Can be incorporated into normal routine

- Fairly straight forward method

- Can provide accurate and detailed information

- Resource efficient - no need to collect the data as it already exists

- Easily accessible, and can be inexpensive

- Can lead to information sharing opportunities

Disadvantages

- Documents are always produced for a certain purpose - can be difficult to relate to your evaluation

- Can be time consuming to analyse large quantities of data

- No guarantees of the quality of data - "the method is as good as the documents"

- Doesn't answer the "why" questions

Validity and Reliability

- No guarantees about the creditability and authenticity of the records and/or the documentation

Resource Requirements and Cost

- Access to documentation

- Time for analysis

Analysis

- Principles of Qualitative or Quantitative Analysis, depending on the documents

Additional Information

- To be used with other tests (Triangulation)

- Evaluators need to know exactly what they are looking for - important to have well established evaluation questions

Tips for a Record / Document Analysis

Internal and external documents are a source of potentially valuable data for your evaluation. These can include project reports, activity schedules, funding proposals, participant records, literature, etc. Such materials enable the evaluator to learn about the history, goals and outcomes of a particular project, and also provide clues about important shifts in the development of the initiative or its maturation. All project related reports are particularly helpful in learning how the project originated, how it is organised, what it claims to do, how it intends to reach its objectives and the nature of its target population.

Always bear in mind that written documents do not necessarily provide comprehensive or correct answers to specific problems, as they may contain errors, omissions, or exaggerations. They are simply one form of evidence, and should be used carefully and in connection with other types of data[7].

[7] Adapted from Kellogg Foundation Evaluation Handbook

Method E – Questionnaires

Description

- Questionnaires can be used to collect quantitative and qualitative data and to measure attitudes, values, personal experiences and behaviour

- Open ended questions add depth to quantitative data and help to further explore the reasons behind statistics

Application

- Could be used as part of any evaluation process of most of the initiatives

- Questionnaires suit both small (CYP projects) and large (HFRA) scale initiatives and can be used to measure both soft and hard outcomes.

Advantages

- Quantitative questionnaires are easy to analyse

- Cheaper and less time consuming than interviews

- Easy to understand

- Provide information with the potential to be quantified

Disadvantages

- Qualitative questionnaires can be time consuming to analyse

- Quantitative questionnaires are not very flexible

- Question wording can cause problems

- Not suitable for illiterate and non-English speaking participants

Validity and Reliability

- Depends on the skills of those designing the questionnaire

- Need to pay attention to design to avoid validity and reliability problems

Resource Requirements and Cost

- Time to develop, administer and analyse the questionnaires

Analysis

- Principles of quantitative or qualitative analysis, depending on the questions

Additional Information

- Good for pre and post tests

Method E – Questionnaires (continued)

How to Design a Questionnaire[8]?

When designing a questionnaire, it is important to pay extra attention to the questionnaire design. In order to obtain accurate and relevant information, you have to give some thought to what questions you ask, how you ask them, the order you ask them in, and the general layout of the questionnaire.

Format of the Questionnaire

Introduction

It seems a good idea to have either a personalised covering letter or at least an introduction explaining briefly the purpose of the questionnaire, the importance of the respondents' participation, who is responsible for the questionnaire, and a statement guaranteeing confidentiality.

Arranging the questions

The order of the questions is also important. Some general rules are:

- Go from general to particular.
- Go from easy to difficult.
- Go from factual to abstract.
- Start with closed format questions.
- Start with questions relevant to the main subject.
- Do not start with demographic and personal questions.

Wording of Individual Questions

The way questions are phrased is important and there are some general rules for constructing good questions in a questionnaire.

Use short and simple sentences

Short, simple sentences are generally less confusing and ambiguous than long, complex ones. As a rule of thumb, most sentences should contain one or two clauses. Sentences with more than three clauses should be rephrased.

Ask for only one piece of information at a time

For example, "Please rate the presentation in terms of its content and presentation" asks for two pieces of information at the same time. It should be divided into two parts: "Please rate the presentation in terms of (a) its content, (b) its presentation."

Avoid negatives if possible

Negatives should be used only sparingly. For example, instead of asking students whether they agree with the statement, "Small group teaching should not be abolished," the statement should be rephrased as, "Small group teaching should continue." Double negatives should always be avoided.

[8] Leung WC "How to conduct a survey"

Ask precise questions

Questions may be ambiguous because a word or term may have a different meaning. For example, if you ask the participants to rate their feelings towards the emergency services, this term could mean different things to different people. For some it could mean the Police and Ambulance Services, and for others it could mean only the Fire Service.

Another source of ambiguity is a failure to specify a frame of reference. For example, in the question, "How often did you attend the boxing club?" the time reference is missing. It might be rephrased as, "How many times have you attended the boxing club within the last two weeks?"

Format of responses

The responses can be in open or closed formats. In an open ended question, the respondents can formulate their own answers. In closed format, respondents choose between several given options. It is possible to use a mixture of the two formats - for example, give a list of options, with the final option of "other" followed by a space for respondents to fill in other alternatives.

Sample open ended questions:

- Why did you first come to the boxing club?

- Have your reasons for attending changed in any way since joining the boxing club?

- What have you learnt at the boxing club, in addition to boxing?

Sample closed questions:

- Were you satisfied with the course delivery? (delete as appropriate)

  Yes / No / I don't know

- How often do you come to the club? (tick one)

  ☐ Twice a week
  ☐ Once a week
  ☐ Less than once a week

Sample scales:

- On a scale of 1-5, to what extent did the young person show willingness to undertake tasks and challenges?
  (1 being no interest, 5 being actively volunteering for activities/tasks)

- On a scale of 1-5, to what extent did the young person show understanding of the value and benefit of team working?
  (1 being no understanding, 5 being actively participating in team work, and encouraging others)

Method F – Cost Analysis

Different approaches to using cost analysis in evaluation[9]:

Cost Benefit – evaluates the project in terms of costs. It measures both the project costs and the results in monetary terms. This means that the results or benefits of the project must be translated into a monetary value. Social Return on Investment can help you to translate the non-tangible benefits of a project into a monetary value. A step by step guide is available from:

http://www.cabinetoffice.gov.uk/third_sector/news/news_stories/090512_sroi.aspx

Cost Effectiveness – evaluates how the desired benefits can be achieved with the least amount of resources. Benefits are expressed only in terms of the impacts or outcomes themselves (they are not given a monetary value). Interpretation of this type of analysis requires stakeholders to decide if the benefit received is worth the cost of the project, or if there are other less expensive projects that would have similar or more beneficial results.

Cost Allocation – focuses on the total costs of the project and compares it against the planned resources. Cost allocation is a simpler concept than either cost-benefit analysis or cost-effectiveness analysis. Cost benefit and cost effectiveness are difficult to analyse, and can show very inaccurate results. At the project level, it basically means setting up budgeting and accounting systems in a way that allows project managers to determine a unit cost or cost per unit of service.

What Cost Analyses can tell you:

- Cost analyses can provide an estimate of what a project costs

- Cost analyses may improve understanding of project operation, and tell what parts of intervention are most cost-effective

- Cost analyses may reveal unexpected costs

What cost analysis cannot tell you:

- Whether or not the project is having a significant net effect on the desired outcomes

- Whether the least expensive alternative is always the best alternative

[9] M. Sewell and N. Marczak "Using Cost Analysis in Evaluation"

3.7 Making Sense of Findings

Once you have gathered all the data, you then
need to analyse and present it in a form that
everyone can understand. The evaluation
findings should inform the decision making
process, therefore the findings have to be in an
easy to understand format. Don't be tempted to
make the findings too simple though, i.e. one
view, one percentage or one cost/benefit ratio
only presents one point of view, and cannot
explain the complexity of a whole project. Always
present both positive and negative findings, and
remember research offers probabilities, not
absolutes - all qualitative and quantitative data
contains varying degrees of error[10].

An evaluator rarely has all the information to
make informed decisions about the future of the
project. Therefore it is important to involve all the
relevant stakeholders in this process. Patton's
framework for reviewing data[11] is particularly
useful for self-evaluations:

1. Description and Analysis

Describing and analysing data findings involves
organising raw data into a form that reveals
basic patterns. The evaluator presents, in user
friendly fashion, the factual finding as revealed in
actual data.

In case your data does not come in numerical
form, the guidance below will help you to put
your qualitative data in a more easy to report
form.

Thematic Data Analysis
Categorise the paragraphs or sentences
according to the themes. Themes are recurrent
topics that are present in the data and could be
describing the same theme but just using
different words[12].

---

[10] Michael Quinn Patton "Utilisation Focused Evaluation"
[11] Michael Quinn Patton "Utilisation Focused Evaluation"
[12] Adapted from Carla Palmer - Derbyshire FRS

Process : name or label the theme, define the theme and then identify data relevant to the theme, e.g. quotations or observations. See below for example.

Lack of effective communication	Low cohesion in group	Role confusion
I haven't been told what happens next and it makes me feel uncertain	Committee is made up of a lot of little Gods… rife with personality conflicts	Half the time you don't know what you're suppose to be doing
The message wasn't passed on to me, and I wasn't aware of the tasks	I am not sure how people see me in the group, whether they think I doing a good job or not	Don't really know sometimes what they expect me to do or how they want it done

Formalised Content Analysis

This is where you count the instance of event or word. It is a form which allows you to change qualitative into quantitative information which can be easier to analysis e.g. you have a paragraph that mentions a negative regarding the scheme that the cost is too expensive three times then this would have a count of three. You would then look for the balance, the positive which could be someone mentions the cost was justified then you could do a comparison possibly 3 negative to 1 positive.

2. Interpretation

What do the results mean? What's the significance of the findings? What are possible explanations of the results? Interpretations go beyond the data to add context, determine meaning and tease out substantive significance based on deduction or inference.

3. Judgement

Values are added to analysis and interpretations. Determining merit or worth means resolving to what extent and in what ways the results are positive or negative. What is good or bad, desirable or undesirable, in the outcomes? Have standards of desirability been met?

Stakeholders need to decide the level at which the project is considered effective/good use of resources. For example if 55% of participants have learnt something during the project, could that be considered as high or low? Ineffective or effective use of resources? The level of impact the project aims to make has to be decided with the stakeholders.

4. Recommendations

The final step (if agreed to be undertaken) adds action to analysis, interpretation, and judgement. What should be done? What are the action implications of the findings? Only recommendations that follow from and are grounded in the data ought to be formulated.

Part 4 – Report On Your Findings

4.1 Reporting Evaluation Findings

Evaluation is only worthwhile if the findings are reported and utilised in an appropriate manner. The recommended structure for a report is:

1. **Executive Summary**

   Summarise the main points from the evaluation including findings and key recommendations.

2. **Introduction to Project**

   Give a full description of the initiative including the overall goal and the context in which it took place.

3. **Evaluation Methods**

   Give a full description of evaluation methods, the sample and explain how the data was gathered and analysed. Also, provide an explanation of the limitations of the evaluation.

4. **Findings**

   This section presents your results. Present quantitative data as tables, pie charts or graphs where appropriate. The interpretation of qualitative data will be presented in textual form. The use of quotes can illustrate the basis of your interpretation and provide valuable insights. The confidentiality of the participants should be respected.

5. **Summary**

   Reflect on the broader lessons from your evaluation, and conclude the study.

6. **Recommendations**

   Make recommendations on the basis of the findings. Recommendations identify what was successful and should be maintained or expanded, and where changes to initiatives seem necessary.

7. **Acknowledgments**

   Thank people who were involved in your evaluation.

8. **References**

   If you have you have used any literature in your evaluation, list any references in a style which enables a reader to follow up on them and is consistent.

9. **Appendices**

   You can include copies of your data collection tools, statistical data etc. that would interrupt the flow of the main report.

4.2 Sharing Evaluation Findings

To make full use of the evaluation findings, make plans for the following activities:

- Share findings with the rest of the GMFRS and partner agencies

- Inform funding sources about the accomplishments

- Use findings to make decisions about the future of the project

- Continue to use evaluation to improve the project and to monitor outcomes

- Continue to share information with stakeholders

- Assess project fit with other communities

Part Four

Further Reading

If you require more information about different types of evaluations, evaluation processes and/or data collections methods, some useful resources are listed below:

Evaluation

Patton, M. "Utilization Focused Evaluation" 3rd ed., 1997, California, Sage

Robson, C. "Small Scale Evaluation" 2000, London, Sage

Rossi et al, "Evaluation: a Systematic Approach" 7th ed., 2004, California, Sage

W.K. Kellogg Foundation - http://www.wkkf.org

Evaluation Trust - http://www.evaluationtrust.org/evaluation/evaluate

Research Methods

Robson, C. "Real Word Research" 2nd ed., 2002, Oxford, Blackwell

Bryman, A. "Social Research Methods", 2008, Oxford, Oxford University Press
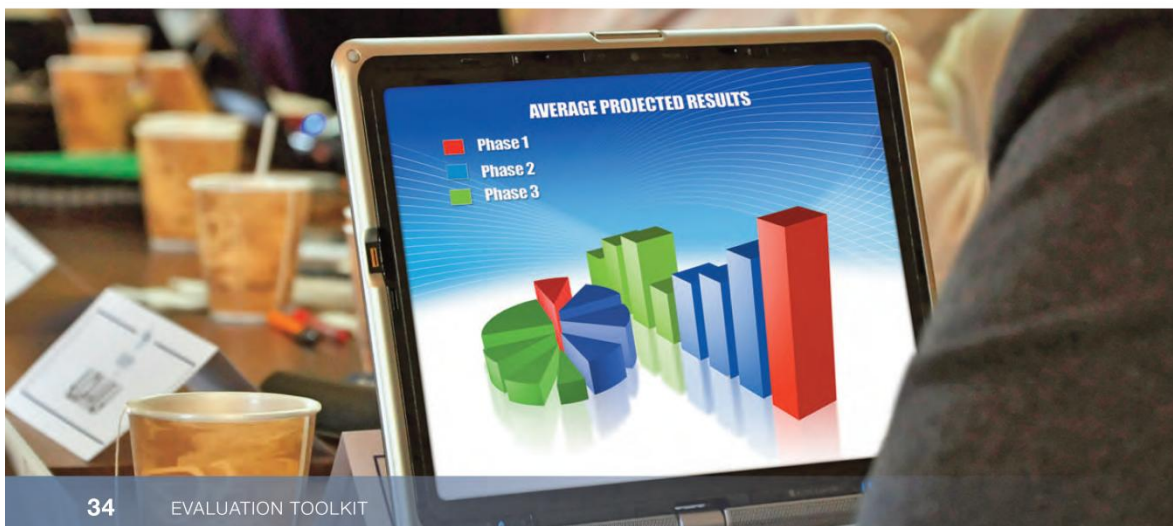
Appendices

Appendix A – Evaluation Summary

Project Planning Stage	Specify aims and objectives of planned activities	☐
Part 1	Decide to evaluate and develop an overall evaluation strategy, including purpose, timeline and budget for evaluation	☐

Before Evaluation	Develop a detailed evaluation plan	☐
Part 2, Part 3	Inform all necessary parties of the evaluation	☐
	Choose, develop and test data collection tools	☐

During Evaluation	Collect Data	☐

After Evaluation	Analyse data and interpret results	☐
Part 4	Write an evaluation report	☐
	Share findings with the rest of the organisation and stakeholders	☐
	Inform funding sources about your evaluation	☐
	Use findings to improve project	☐

Appendix B – Project Logic Model[13]

Project Aim:						
	Inputs	**Outputs**		**Outcomes – Impact**		
	What we invest:	What we do:	Who we reach:	Short term:	Medium term:	Long term:
Assumptions: (the belief you have about the project / the way you think the project will work)				**External Influencing Factors:**		

[13] University of Wisconsin Program Development and Evaluation 2002

Appendix C – Evaluation Plan

1. Evaluation Focus Area	2. Evaluation Question	3. Indicators	4. Method

Appendix D – Timetable for Evaluation

[Insert name of initiative]	Year												
	Task:	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec

235

Appendix E – Key Word Definitions

Aim

The ultimate goal(s) of the project. Answers the questions: "What is the project going to achieve?"

Anonymity

The participant should not be identifiable from any published materials. Participant's personal details, including name, should not be shared with a third party.

Control Group

A group of people that is matched as closely as possible with participants, but who are not participating in the project, can be used to validate the changes that occur in the participant's group as a consequence of the initiative.

Data Collection Tools / Research Methods

The tools that allow you carry out your research and to examine your evaluation topic, for example questionnaires, interviews, etc.

Hard Outcomes

Hard outcomes refer to the quantitative outcomes the initiatives aim to achieve. Hard outcomes often appear as numbers, statistics and percentages.

Objective

Objectives are the ways to achieve your aim(s), the activities needed to carry out to complete the task. Answers the questions: "How are you going to achieve the aim?"

Primary Research

Primary research (also called field research) involves the collection of data that does not already exist.

Qualitative Data

Non-numerical, in depth data, that investigates the why and how questions.

Quantitative Data

Numerical data or data that can be converted into numbers, and is measured as how many, how long how much, etc.

Reliability

Reliability is the extent to which a data collection tool yields the same result on repeated trials. In other words, will the tool provide the same answers even if it is administered at different times or by different people.

Sampling

Sampling is concerned with the number of participants taking part in the research.

Secondary Research

Data that has already been collected for another purpose.

Soft Outcomes

Soft outcomes are qualitative in nature, and difficult to quantify. Soft outcomes are often expressed in words and refer to non-tangible issues such as behaviour, awareness and attitudes.

Structured Interviews

The interviewer asks predetermined questions, exactly as they are written, in the same sequence and using the same style.

Triangulation

Triangulation refers to the use of multiple research methods. Evaluation designs should incorporate both qualitative and quantitative data collection methods whenever possible.
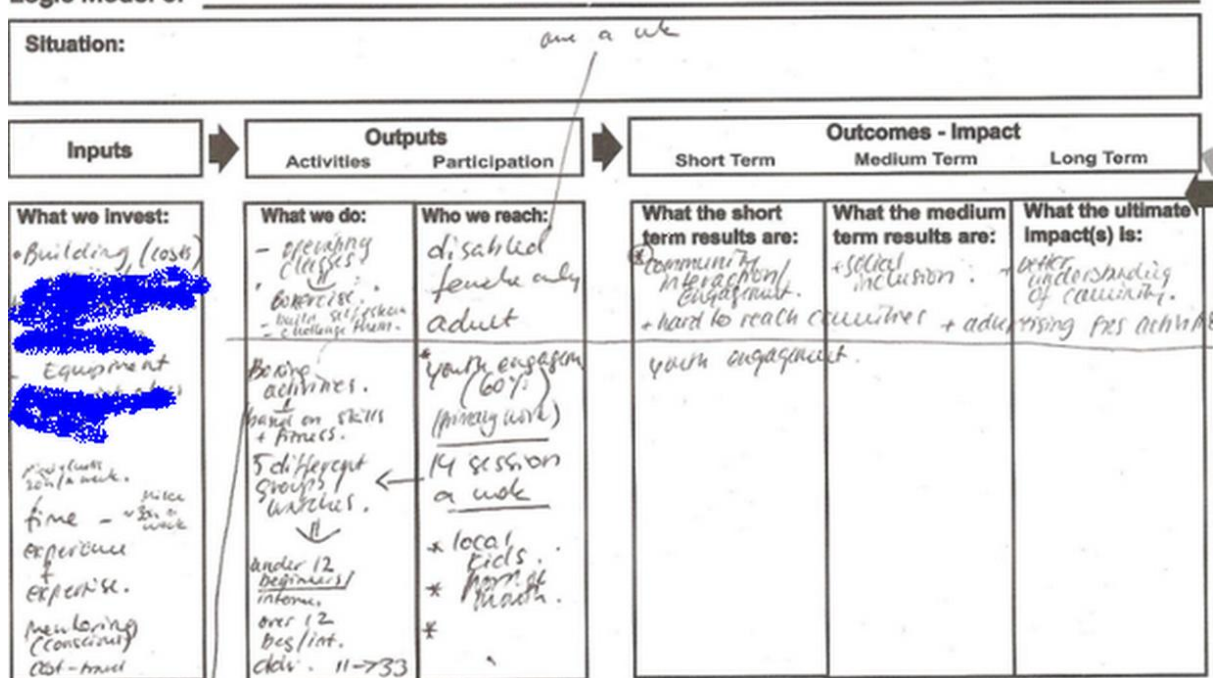
Validity

Validity refers to the degree to which a study accurately reflects the specific concept that the evaluator is attempting to measure. i.e. "Does the measure really measure what it was set to measure?"

Appendix E – Logic Model Before and After

1. Logic Model before

2. Logic Model after

Project Aim : To engage with local youth to prevent them from attacking firefighters and setting fires					
Inputs	Outputs - Activities	Outputs - Participation	Outcomes Short Term	Outcomes Medium Term	Outcomes Long Term
Building. Equipment. Time.	Teach boxing to 5 different groups, based on their skills and fitness levels. Educate young people about discipline, respect and courage to change.	Local Youths.	Community engagement. Young People - awareness of Fire Service activities, and positive attitude towards fire fighters. Young people - awareness of good behaviour.	Young people - increased self esteem, opinions of FRS changed, and behaviour improved. FRS - better understanding of community needs.	Less attacks on fire fighters. Less fire crime. Safer communities.
Assumption: Boxing has a certain kudos, it increases self esteem, and it can increase the respect that an individual gets from their peer group. If these young people have an influence in a gang or group situation and they have an affinity to the fire service then, we are in a win-win situation where we can help to form citizens of the future whilst having ambassadors for the Fire Service on the streets of Moss Side.				**External Factors:** Gangs GMFRS support Partner agencies Trainers availability	

Appendix F – Evaluation Timetables

1. The initial timetable:

Action	June					July				August					September			
week c/o	01-Jun	08-Jun	15-Jun	22-Jun	29-Jun	06-Jul	13-Jul	20-Jul	27-Jul	03-Aug	10-Aug	17-Aug	24-Aug	31-Aug	07-Sep	14-Sep	21-Sep	28-Sep
Training					▓													
Interviews with Participants and Parents					▓	▓	▓											
Interviews with Trainers					▓													
Interviews with Partners					▓													
Document Analysis								▓	▓									
Data Analysis											▓	▓	▓	▓				
Report Writing											▓	▓	▓	▓				

2. Amended timetable (due to participants failure to complete tasks on time)

Action	June					July				August					September			
week c/o	01-Jun	08-Jun	15-Jun	22-Jun	29-Jun	06-Jul	13-Jul	20-Jul	27-Jul	03-Aug	10-Aug	17-Aug	24-Aug	31-Aug	07-Sep	14-Sep	21-Sep	28-Sep
Training					▓													
Interviews with Participants and Parents							▓	▓	▓									
Interviews with Trainers						▓				▓								
Interviews with Partners						▓		▟R	▟R	▓								
Postcode Analysis								▟R	▟R									
Document Analysis (SPSS)								▓	▓									
Data Analysis - interviews, post codes,											▓	▓		▓				
Report Writing														▓	▓			

Appendix G - The SARA Model

SARA model (Scanning, Analysis, Response and Assessment) is a commonly used problem-solving method associated with Problem-Oriented Policing. POP is an approach to policing in which activities are subject to a thorough examination in order to discover new and more effective ways of dealing with them (Goldstein 2001). The SARA model contains the following elements (Center for Problem-Orientated Policing 2012):

Scanning:

- Identifying recurring problems of concern to the public and the police.

- Identifying the consequences of the problem for the community and the police.

- Prioritizing those problems.

- Developing broad goals.

- Confirming that the problems exist.

- Determining how frequently the problem occurs and how long it has been taking place.

- Selecting problems for closer examination.

Analysis:

- Identifying and understanding the events and conditions that precede and accompany the problem.

- Identifying relevant data to be collected.

- Researching what is known about the problem type.

- Taking inventory of how the problem is currently addressed and the strengths and limitations of the current response.

- Narrowing the scope of the problem as specifically as possible.

- Identifying a variety of resources that may be of assistance in developing a deeper understanding of the problem.

- Developing a working hypothesis about why the problem is occurring.

Response:

- Brainstorming for new interventions.

- Searching for what other communities with similar problems have done.

- Choosing among the alternative interventions.

- Outlining a response plan and identifying responsible parties.

- Stating the specific objectives for the response plan.

- Carrying out the planned activities.

Assessment:

- Determining whether the plan was implemented (a process evaluation).

- Collecting pre– and post–response qualitative and quantitative data.

- Determining whether broad goals and specific objectives were attained.

- Identifying any new strategies needed to augment the original plan.

- Conducting ongoing assessment to ensure continued effectiveness.

Appendix H – Evaluation Toolkit Workshop Questions

1. Is the toolkit…

    a) Flexible – could it be used for any project?

    b) Easy to use - Would you be able to carry out evaluations using the toolkit?

    c) Clarity – Are all the different stages of evaluation explained in a clear manner?

    d) Sufficient amount of details – Does the toolkit cover all the different stages and

    processes in enough detail?

2. Is there anything you would like to see added or changed?

3. In addition to the toolkit, should the organisation provide any further assistance with

    evaluation?

Appendix I - Questions for Toolkit User Interviews

Process:

1. a) What do you think is the purpose of evaluation (in this organisation)?

   b) Has your view changed since taking part in the evaluation trial?

2. How did you find the evaluation process? Would you do something differently?

3. Now that you have got experience with evaluation, what do you think are the advantages and disadvantages of implementing evaluation?

4. Has the evaluation had any impact on the initiative?

5. How do you think evaluation will affect the initiatives in the future

Toolkit feedback:

6. Is it fit for purpose?

   a) Flexible – could it be used for any project?

   b) Easy to use - Would you be able to carry out evaluations using the toolkit?

   c) Clarity – Are all the different stages of evaluation explained in a clear manner?

   d) Enough details – Does the toolkit cover all the different stages and processes in enough detail?

7. In addition to the toolkit – would you need any other assistance to evaluate a project?

8. Do you think the organisation will employ the toolkit / Do you think the toolkit will help evaluation to become an embedded process?

References

Adams, J. & Dickinson, P. (2010). Workforce Evaluation Training to Build Capability in the Community and Public Health. *American Journal of Evaluation* 31(42), 421-433.

Adler, P.A. & Adler, P. (1994). 'Observational Techniques'. In Denzin, Norman K. & Yvonna S. Lincoln, (eds.) (1994). *Handbook of Qualitative Research*. (pp. 377-392). Thousand Oaks, CA: Sage Publications.

Aim Higher Greater Merseyside (2006). *Evaluation Toolkit.* Retrieved 29 January 2009 from: www.aimhigher.ac.uk.

Alkin, M. (1972). *A Classification Scheme for Objectives-based Evaluation Systems.* Los Angeles, Evaluation Theory Program, UCLA Graduate School of Education.

Alkin, M. (2004). *Evaluation Roots: Tracing Theorists' Views and Influences*. California, Sage.

Annabel Jackson Associates (2004). *Evaluation Toolkit for the Voluntary and Community Arts*. Retrieved 4 February 2009 from: www.artscouncil-ni.org.

Argyris, C. & Schön, D. (1974) *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass.

Arnold, M. (2006). Developing Evaluation Capacity in Extension 4-H Field Faculty: A Framework for Success. *American Journal of Evaluation*, 27(2), 257-269.

Aronson, J. (1994). A Pragmatic View of Thematic Analysis. *The Qualitative Report*, 2(1), Spring.

Atkins S. & Murphy K. (1994). Reflective Practice. *Nursing Standard*, 8(39), 49-56.

Atkinson & Coffey (2004) in Bryman, A. (2008). *Social Research Methods*. (pp.514-532) Oxford, Oxford University Press.

Bain, G., Lyons, M. & Young, A. (2002). *The Future of the Fire Service: Reducing Risk, Saving Lives: the Independent Review of the Fire Service December 2002*. London, ODPM.

Baizerman, M., Compton, D.W. & Stockdill, S.H. (2002). New Directions for ECB. *New Directions For Evaluation,* 93, Spring.

Barnette, J. & Wallis, A. (2003). 'Helping Evaluators Swim with the Current: Training Evaluators to Support Mainstreaming'. In Barnette, J. & Sanders, J. (Eds.), Mainstreaming Evaluation, *New Directions for Evaluation*, 99, 51-61. Jossey-Bass.

Baskerville, R.L. (1999). Investigating Information Systems with Action Research. *Communications of the Association for Information Systems*, 2(4).

Betzner, A.E. (2008). *Pragmatic and Dialectic Mixed Method Approaches: An Empirical Comparison,* PhD Thesis, University of Minnesota. Retrieved from http://purl.umn.edu/46961

Blaikie, N. (2007). *Designing Social Research.* Cambridge, Polity.

Borton, T. (1970). 'Framework Guiding Reflective Activities' in Jasper, M. (2003). *Beginning Reflective Practice.* Cheltenham: Nelson thorn

Boud, D., Keogh, R. & Walker, D. (1985). *Reflection: Turning Experience into Learning.* London, Kogan Page.

Boulmetis, J. & Dutwin, P. (2005). *The ABCs of Evaluation.* San Francisco, Jossey-Bass

Boyle, R. & Lemaire, D. (edt.) (1999). *Building Effective Evaluation Capacity: Lessons from Practice.* London, Transaction Publishers.

Brannen, J. (1995). *Mixing Methods: Qualitative and Quantitative Research.* Aldershot, Avebury.

Braun, V. & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology,* 3(2), 77-101.

Bryman, A. (2006). Integrating Quantitative and Qualitative Research: How Is It Done?. *Qualitative Research,* 6(1), 97-114.

Bryman, A. (2008). *Social Research Methods*. 3rd ed. Oxford, Oxford University Press.

Byrne, M. (2001). Data Analysis Strategies for Qualitative Research. *AORN*. 74 (6), 904-5.

Caracelli, V.J. (2000). Evaluation Use at the Threshold of the Twenty-First Century. *New Directions for Evaluation,* 2000(88), 99–111.

Catterall, M. & Maclaran. P. (1997). Focus Group Data and Qualitative Analysis Programs: Coding the Moving Picture as well as the Snapshots. *Sociological Research Online*, 2(1), 53-61.

Center for Problem-Oriented Policing. *The SARA Model*. Retrieved 28 May 2010 from: http://www.popcenter.org/about/?p=whatiscpop

Centre for Health Promotion (2005). *Evaluating Comprehensive Workplace Health Promotion Info-Pack*. Retrieved 5 February 2009 from: www.thcu.ca/workplace/documents/EvaluationInfoPackFinalWeb.pdf

Cleland, D.I. (1999). *Project Management: Strategic Design and Implementation*. London, McGraw-Hill.

Cohen C. (2006). Evaluation Learning Circles: A Sole Proprietor's Evaluation Capacity-Building Strategy. *New Directions for Evaluation*, 111, 85-93.

Compton, D. W. (2009). Managing Studies Versus Managing for Evaluation Capacity Building. *New Directions for Evaluation,* 121, 55–69.

Cousins, B.J. & Leithwood K.A. (1986). Current Empirical Research on Evaluation Utilization. *Review of Educational Research*, 56(3), 331-364.

Cousins, J., Goh, S. & Clark, L. (2004). Integrating Evaluative Inquiry into the Organisational Culture: A Review and Synthesis of the Knowledge Base. *The Canadian Journal of Program Evaluation*, 19(2), 99–141.

Creswell, J.W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Method Approaches*. London, Sage.

Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London, Sage.

Davidson , E. J. (2001). The meta-learning organization: A model and methodology for evaluating organizational learning capacity. *Dissertation Abstracts International,* 62(5).

Davidson, E.J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks: Sage.

Davies (1999) in McDonald B., Rogers, P. & Kefford, B. (2003). Teaching People to Fish? Building the Evaluation Capability of Public Sector Organizations. *Evaluation* 9(1): 9–29.

Dawson, K. (2009). *Introduction to Research Methods: A Practical Guide for Anyone Undertaking a Research Project*. Oxford, How to Books Ltd.

Denzin, N.K. & Lincoln, Y.S. (2000). 'The Discipline and Practice of Qualitative Research'. in Denzin, N.K. & Lincoln, Y.S. (Edt). *The Handbook of Qualitative Research*. (pp.1-28). 2nd ed. Sage, London.

Department of Health and Human Services (2004). *How to Evaluate Safety and Health Changes in the Workplace*. Retrieved 5 February 2009 from:
http://www.cdc.gov/niosh/docs/2004-135/downloads/2004-135.pdf

Dewson, S., Eccles, J., Tackey, N.D. & Jackson, A. (2000). *Guide to Measuring Soft Outcomes and Distance Travelled*. Brighton, the Institute for Employment Studies.

Dick, B. (1993). You Want to Do an Action Research Thesis? *Action Research Resources*, Retrieved 18 August 2011 from:
http://www.scu.edu.au/schools/gcm/ar/art/arthesis.html

Dick, B. (2002). *Questions for Critical Reflection [On line]*. Retrieved 12 January 2012 from: http://www.uq.net.au/action_research/arp/reflques.html

Duignan, P. (2003). Mainstreaming evaluation or building evaluation capability? Three key elements. *New Directions for Evaluation* 99, 7-21.

Evaluation Assistance Center (1995). *Evaluation Handbook*. Retrieved 25 February 2009 from: http://www.ncela.gwu.edu/files/rcd/BE020502/Evaluation_Handbook.pdf

Fetterman, D. & Wandersman, A. (2007). Empowerment Evaluation: Yesterday, Today, and Tomorrow. *American Journal of Evaluation*, 28(2), 179-198.

Fire Service Act 1947. Retrieved 18 January 2012 from: http://www.legislation.gov.uk/ukpga/1947/41/pdfs/ukpga_19470041_en.pdf

Fire Service Act 2004. Retrieved 18 January 2012 from: http://www.legislation.gov.uk/ukpga/2004/21/pdfs/ukpga_20040021_en.pdf

Fontana, A., & Frey, J. H. (1994). 'Interviewing: The Art of Science'. In N. Denzin & Y. Lincoln (Eds.), *Handbook of Qualitative Research,* (pp. 361-377). London: Sage.

Forss K., Kruse, S-E., Taut, S & Tenden, E. (2006). Chasing a Ghost? An Essay on Participatory Evaluation and Capacity Development. *Evaluation,* 12(1), 128–144.

Franzosi, R.P. (2004) in Hardy, M. & Bryman, A. (2004). *Handbook of Data Analysis*. London, Sage.

Gibbs, A. (1997). Focus groups. *Social Research Update*. [On-line], Issue Nineteen. University of Surrey. Retrieved 15 September 2011 from: http://www.soc.surrey.ac.uk/sru/SRU19.html.

Gibbs, D., Napp, D., Jolly, D., Westover, B. & Uhl, G. (2002). Increasing Evaluation Capacity Within Community-based HIV Prevention Programs. *Evaluation and Program Planning,* 25, 261–269.

Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Oxford: Further Education Unit, Oxford Polytechnic.

GMFRS (2011) *Corporate Plan 2011-2014*, Greater Manchester Fire and Rescue Service, Retrieved 18 January 2012 from:

http://www.manchesterfire.gov.uk/media/111806/corporate_plan_2011.pdf

Goldstein, H. (2001). *History of Problem-Oriented Policing*. (On-line). Retrieved 28 May 2010 from: http://www.popcenter.org/about/?p=whatiscpop

Greater Manchester Fire and Rescue Authority (2004). *Integrated Risk Management Plan 2004-2007*, Greater Manchester Fire and Rescue Service, Retrieved 19 March 2012 from:

http://www.manchesterfire.gov.uk/about_us/policies_and_publications/irmp.aspx

Grudens-Schuck, N. (2003). The Rigidity and Comfort of Habits: A Cultural and Philosophical Analysis of the Ups and Downs of Mainstreaming Evaluation. *New Directions for Evaluation*, 99, 23-32.

Guba, E.G. (1990). *The Paradigm Dialog*. California, Newbury Park.

Guba, E.G. & Lincoln, Y.S. (2003). 'Competing Paradigms in Social Research: Theories and Issues'. in Hesse-Biber, S & Leavy, P. (Edt). *Approaches to Qualitative Research: A Reader on Theory and Practice*. (pp.17-38). Oxford University Press, New York.

Harvey, J. (ed.) (1998). *The Evaluation Cookbook*. Edinburgh: Heriot-Watt University.

Hay, K. (2010). Evaluation Field Building in South Asia: Reflections, Anecdotes, and Questions. *American Journal of Evaluation*, 31(2), 222-231.

Health and Safety Executive (1998). HSG65 Model. Retrieved 4 November, 2009, from http://www.hse.gov.uk/pubns/indg275.pdf

Hesse-Biber, S & Leavy, P. (2003). 'Distinguishing Qualitative Research'. in Hesse-Biber, S & Leavy, P. (Edt). *Approaches to Qualitative Research: A Reader on Theory and Practice*. (pp.1-16) Oxford University Press, New York

Hodder, I. (1994). 'The Interpretation of Documents and Material Culture'. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 393-402). Thousand Oaks, CA: SAGE Publications.

Howitt, D. & Cramer, D. (2008). *Introduction to Research Methods in Psychology*. Essex, Pearson Higher Education.

Huffman, D., *Lawrenz, F., Thomas, K. & Clarkson, L.* (2006). Collaborative Evaluation Communities in Urban Schools: A Model of Evaluation Capacity Building for STEM Education. *New Directions for Evaluation,* 109, Spring.

Huffman, D., Thomas, K. & Lawrenz, F. (2008). A Collaborative Immersion Approach to Evaluation Capacity Building. *The American Journal of Evaluation,* 29:3, 358-368.

Institute for Work and Health, *Guide to Evaluating the Effectiveness of Strategies for Preventing Work Injuries,* (2001). Retrieved 25 February from:

http://www.cdc.gov/niosh/docs/2001-119/pdfs/2001-119.pdf

Johns, C. (1994). *Model of Reflection*. Retrieved 12 2012 from:

http://www.ahot.utu.fi/yliopistot/materiaali/index/using_a_model_of_reflection.pdf

Johnson, R.B. & Onwuegbuzie, A.J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher,* 33(7), 14-26.

Khan, M. (1998). Evaluation Capacity Building - An Overview of Current Status, Issues and Options. *Evaluation,* 4(3), 310–328.

King J. A., & Volkov, B (2005) in Taylor-Powell E. & Boyd H. (2008). Evaluation Capacity Building in Complex Organizations. *New Directions For Evaluation,* 120, Winter.

King, J. (2002). Building the Evaluation Capacity of a School District. *New Directions For Evaluation,* 93, 63-80.

Kolb, D.A. (1984). *Experiential Learning Experience as a Source of Learning and Development.* New Jersey: Prentice Hall

KTP Grant Application and Proposal Form (2007), Salford, University of Salford / KTP

Kuzim A. (2009). *Evaluation Capacity Building Strategy: Towards a Mature Profession.* Paper presented at the Conference on National Evaluation Capacity, Casablanca, Morocco, December 15-17, 2009.

Lennie, J. (2005). An Evaluation Capacity-Building Process for Sustainable Community IT Initiatives. *Evaluation,* 11:4, 390–414.

Levin, M. & Greenwood, D. (2001). 'Pragmatic Action Research and the Struggle to Transform Universities into Learning Communities'. in Reason, P. & Bradbury, H. (eds.) *Action Research - Participative Inquiry and Practice*. (pp. 103-113). Sage, London.

Leviton, L. C. & Hughes, E. F. (1981). Research on the Utilization of Evaluations: A Review and Synthesis. *Evaluation Review,* 5(4), 525–548.

Lincoln, Y.S. (1989). *Fourth Generation Evaluation.* California, Sage.

Lincoln, Y.S. & Guba, E.G. (1985) in Denzin, N.K. & Lincoln, Y.S. (Edt) (1994). *The Handbook of Qualitative Research*. 2nd ed. London, Sage.

Lock, D. (2007). *Project Management.* (9th ed.) Hampshire, Gower Publishing, Ltd.

Longman, A. & Mullins, J. (2005). *The Rational Project Manager: A Thinking Team's Guide to Getting Work Done*. Hoboken. New Jersey, John Wiley & Sons Inc.

MacIsaac. (1995). 'Action Research Model', in O'Brien, R. (1998). *An Overview of the Methodological Approach of Action Research*. Faculty of Information Studies, University of Toronto, Retrieved 21 June 2010 from:

http://www.web.net/~robrien/papers/arfinal.html

Marshall (2001). 'Self-Reflective Inquiry Process' in Reason, P. And Bradbury, H. (eds) *Handbook of action research: participative enquiry and practice*. (pp.433-439). London: Sage.

Mason, J. (2002). *Qualitative Researching*. 2nd ed. London, Sage.

Mason, J. (2006). Mixing Methods in a Qualitatively Driven Way. *Qualitative Research*, 6(1), 9–25.

May, T. (2001). *Social Research: Issues, Methods and Processes* (3rd ed). Buckingham, Open University Press.

McDonald B., Rogers, P. & Kefford, B. (2003). Teaching People to Fish? Building the Evaluation Capability of Public Sector Organizations. *Evaluation*, 9(1), 9–29.

McNamara, G., O'Hara, J., Boyle, R. & Sullivan, C. (2009). Developing a Culture of Evaluation in the Irish Public Sector. *Evaluation,* 15(1), 100-112.

McNiff, J. (2000). *Action Research in Organisations*. London, Routledge.

McNiff, J. & Whitehead, J. (2002). *Action Research: Principles and Practice*. 2nd edt., London, Routledge.

McNiff, J. & Whitehead, J. (2006). *All You Need To Know About Action Research*. London, Sage.

Meredith, J. & Mantel, 2006. *Project Management: A Managerial Approach*. 6th ed. Asia: John Wiley & Sons.

Miller, T., Kobayashi, M.M. & Noble, P.M. (2006). Insourcing, Not Capacity Building, a Better Model for Sustained Program Evaluation. *American Journal of Evaluation*, 27(1), 83-94.

Milstein & Cotton (2001) in Naccarella L., Pirkis, J., Kohn, F., Morley, B., Burgess, P. & Blashki, G. (2007). (2007). Evaluators as Educators: Articulating Program Theory and Building Evaluation Capacity. *Evaluation and Program Planning,* 30, 231–236.

Milstein, B., Chapel, T.J., Wetterhall,S.F. & Cotton, D.A. (2002). Building Capacity for Program Evaluation at the Centers for Disease Control and Prevention. *New Directions For Evaluation*, 93, Spring.

Minnett, A. (1999). Internal Evaluation in a Self-Reflective Organization: One Non-Profit Agency's Model. *Evaluation and Program Planning,* 22, 353-362.

Monroe, M., Fleming,M.L., Bowman R.A., Zimmer,J.F., Marcinkowski,T., Washburn, J. & Mitchell N.J. (2005). Evaluators as Educators: Articulating Program Theory and Building Evaluation Capacity. *New Directions For Evaluation,* 108.

Morgan, D.L. (2007). Paradigms Lost and Pragmatism Regained Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research* 1(1), 48-76.

Naccarella L., Pirkis, J., Kohn, F., Morley, B., Burgess, P. & Blashki, G. (2007). Evaluators as Educators: Articulating Program Theory and Building Evaluation Capacity. *Evaluation and Program Planning,* 30, 231–236.

Nu'Man J., King, W., Bhalakia, A. & Criss, S. (2007). A Framework for Building Organizational Capacity Integrating Planning, Monitoring, and Evaluation. *Public Health Management Practice,* January(Suppl), 24–32.

O'Brien, R. (1998). *An Overview of the Methodological Approach of Action Research.* Faculty of Information Studies, University of Toronto, Retrieved 21 June 2010 from: http://www.web.net/~robrien/papers/arfinal.html

Patton, M. (1997). *Utilisation Focused Evaluation.* California, Thousand Oaks.

Patton, M. (2001). *Qualitative Research & Evaluation Methods*. (3rd ed.), London Sage.

Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. London, Sage.

Phillips, C.J., Palfrey C. & Thomas, P. (1994). *Evaluating Health and Social Care*. London, Macmillan.

Plano Clark, V. & Creswell, J. (2008). *The Mixed Methods Reader*. California, Thousand Oaks.

Porteous, N.L., Sheldrick B.J., & Stewart P.J. (1999). Enhancing Managers' Evaluation Capacity: A Case Study From Ontario Public Health. *The Canadian Journal of Program Evaluation,* Special Issue, 137–154.

Posavac, E. J. & Carey, R. G. (1985). *Program Evaluation: Methods and Case Studies*. (2nd ed.), NJ: Prentice-Hall, Englewood Cliffs.

Preskill H. & Boyle S. (2008). A Multidisciplinary Model of Evaluation Capacity Building. *American Journal of Evaluation,* 29, 443-459.

Preskill, H. (2004). 'The transformative power of evaluation: Passion, purpose, and practice'. In, Alkin, M. (Ed.), *Evaluation roots: tracing theorists' views and influences*. Thousand Oaks, CA: Sage.

Reason, P. & Bradbury, H. (Edt) (2001). *Handbook of Action Research*. London, Sage.

Robson, C. (2000). *Small-Scale Evaluation: Principles and Practice*. London, Sage.

Robson, C. (2002). Real World Research. 2nd ed. Oxford, Blackwell.

Rossi, P.H. & Freeman, H.E. (1993). *Evaluation: a Systematic Approach*. London, Sage.

Rudolph, J.W., Taylor, S.S. & Foldy E.G. (2001). 'Collaborative Off-line Reflection: A Way to Develop Skill in Action Science and Action Inquiry'. In P. Reason & H. Bradbury (Eds), *Handbook of action research: Participative inquiry and practice*. (405–412). London: Sage.

Rutman, L. & Mowbray, G. (1983). *Understanding Program Evaluation*. London, Sage, in co-operation with the University Of Michigan School Of Social Work.

Sanders, J.R. (2002). Presidential Address: On Mainstreaming Evaluation, *American Journal of Evaluation,* 23(3), 253–259.

Sarantakos, S. (2005). *Social Research*. Basingstoke, Palgrave Macmillan.

Scott (1990) in Bryman, A. (2008). *Social Research Methods*. 3rd ed. Oxford, Oxford University Press.

Scriven, M. (1967). 'The Methodology of Evaluation'. In Tyler, R. W., Gagné, R. M. & Scriven, M. (Eds.), *Perspectives of Curriculum Evaluation,* 39-83. Chicago, IL: Rand McNally.

Scriven, M. (1991). *Evaluation Thesaurus*. London, Sage.

Sewell, M. (2009). *The Use of Qualitative Interviews in Evaluation*. Cyfernet-Evaluation. Tucson: University of Arizona. Retrieved 18 August 2010 from:

http://ag.arizona.edu/fcs/cyfernet/cyfar/Intervu5.htms

Somekh, B. (2006). Action Research. *British Journal of Educational Technology* 37(6), 990

Sport England, (n.d.). *Monitoring and Evaluation Toolkit*. Retrieved 12 January 2009 from: http://www.sportengland.org/

Stevenson J.F., Florin, P., Scott Mills, D. & Andrade, M. (2002). Building Evaluation Capacity in Human Service Organizations: a Case Study. *Evaluation and Program Planning*, 25, 233–243.

Stockdill, S.H., Baizerman M. & Compton D.W. (2002). Toward a Definition of the ECB Process: A Conversation with the ECB Literature. *New Directions for Evaluation,* 93, 7-26.

Stringer, E.T. (2007). *Action Research*. 3rd ed. California, Sage.

Swepson, P. (1995). *Action research: Understanding its Philosophy San Improve Your Practice*. Retrieved 16 June 20101 from:

http://www.uq.net.au/action_research/arp/philos.html

Tashakkori, A. & Teddlie, C. (2003). *Handbook of Mixed Methods Social and Behavioral Research*. California, Sage.

Taut, S. (2007). Defining evaluation capacity building: Utility considerations. *American Journal of Evaluation,* 28(1), 120.

Taut, S. (2007). Studying Self-Evaluation Capacity Building in a Large International Development Organization. *American Journal of Evaluation,* 28(1), 45-59.

Taut, S. M., & Brauns, D. (2003). Resistance to evaluation: A psychological perspective. *Evaluation,* 9, 247-264.

Taut, S.M. & Alkin, A.C. (2003). Program Staff Perceptions of Barriers to Evaluation Implementation. *American Journal of Evaluation,* 24(2), 213–226.

Taylor-Powell E. & Boyd H. (2008). Evaluation Capacity Building in Complex Organizations. *New Directions For Evaluation,* 120, 55-69.

The World Bank. (2004). *Monitoring and Evaluation: Some Tools, Methods & Approaches*. Washington, DC: The World Bank. Retrieved 27 January 2009 from: http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/4585672-1251481378590/MandE_tools_methods_approaches.pdf

Torres, R. & Preskill, H. (1999). Building Capacity for Organizational Learning Through Evaluative Inquiry. *Evaluation*, 5(1), 42-60.

Trochim, W.M. (2006). *The Research Methods Knowledge Base. 2nd Edition*. Retrieved 24 August 2010 from: http://www.socialresearchmethods.net/kb/>

Vedung, E. (2010). Four Waves of Evaluation Diffusion. *Evaluation* 16(3), 263-277.

W.K. Kellogg Foundation (2004). *Evaluation Handbook*. Retrieved 16 January 2009 from: http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx

Wandersman, A., Flaspohler, P., Ace, A., Ford, L., Imm, P. S. & Chinman, M. J., (2003). PIE à la Mode: Mainstreaming Evaluation and Accountability in Each Program in Every County of a Statewide School Readiness Initiative. *New Directions For Evaluation,* 99, 33-49.

Wandersman, A., Snell-Johns, J, Lentz B.E., Fetterman D.M., Keener D.C., Livet, M…. Flaspohler, P. (2005). 'The Principles of Empowerment Evaluation'. In Fetterman, D. M. and Wandersman, A. (Eds.). *Empowerment Evaluation Principles in Practice* (pp. 27-41). New York: Guilford.

Williams (2001) in McDonald B., Rogers, P. & Kefford, B. (2003). Teaching People to Fish? Building the Evaluation Capability of Public Sector Organizations. *Evaluation* 9(1), 9–29.

Wing K. (2004). Assessing the Effectiveness of Capacity-Building Initiatives: Seven Issues for the Field. *Nonprofit and Voluntary Sector Quarterly* 33(1), 153-160.

Winter, R. & Munn-Giddings, C. (2002). *A Handbook for Action Research in Health and Social Care*. London, Routledge.