Clinical Leadership: can the skills be learned by trainee paediatricians?

Dr Robert Klaber

UCL

MD (Res)

‘I, Dr Robert Klaber, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Abstract

Aim: To explore whether paediatricians in training can develop leadership skills through participating in a specifically designed leadership development initiative.

Methods: A systematic review was conducted to explore the healthcare leadership literature for empirical evidence of different approaches to leadership development. Informed by this review, and conceptualised by key leadership theories, a work-based leadership development initiative was established within a newly formed trainee committee in the London School of Paediatrics. 32 paediatric trainee participants were recruited, and a comparator group came from within the wider trainee population. The trainee group recruited to the leadership initiative were also offered the opportunity to use a multi-source feedback tool to support their learning. The impact of the leadership development initiative was evaluated using a mixed methods approach. A web-based self-assessment questionnaire was designed, validated and then used as a pre- and post-intervention (12-months) quantitative measure. In addition focus groups were run pre- and post-intervention and framework analysis was used to explore the qualitative data.

Results: The systematic review highlighted small areas of empirical evidence around leadership development within medicine. The quantitative analysis demonstrated a significant increase in the participants' self-assessment scores between the pre-intervention questionnaire and the post-intervention questionnaire. Additionally, paired t-tests demonstrated statistically significant increases in scores relating to 'developing networks and building & maintaining relationships' and 'working within teams and encouraging contribution'. The qualitative analysis highlighted the importance of personal development, team-working and understanding the wider context of healthcare systems emerged as important themes.

Conclusions: This work helps to strengthen the understanding that clinical leadership skills and capabilities can be learned by trainee paediatricians. The results are suggestive of the positive impact of the intervention on many of the individuals involved, and despite some methodological weaknesses, this thesis has generated new evidence that can be used to influence the design and planning of other leadership development initiatives in the future.

Table of Contents

Abstract **3**

List of Figures and Tables **10**

Chapter 1 – Background **11**

1.1 Background to wider leadership literature and theories 11

1.1.1 Introduction to the evolution of theories of leadership 11

1.1.2 Trait theory 11

1.1.3 Leadership styles and behaviours 12

1.1.4 Transactional and transformational leadership 13

1.1.5 Healthcare organisations as professional bureaucracies 14

1.1.6 Shared Leadership 15

1.2 Background to clinical leadership and medical leadership 15

1.2.1 What is clinical leadership? 15

1.2.2 Clinical leadership in the NHS 16

1.2.3 Clinical leadership or medical leadership? 18

1.2.4 Medical Leadership Competency Framework (MLCF) 19

1.2.5 Leadership learning within paediatrics 21

1.2.6 Definitions of leadership 22

1.3 Aims and objectives of the thesis 24

1.3.1 Overall aim of the thesis 24

1.3.2 Overall thesis objectives 24

1.3.3 Orientation of chapters 24

1.4 Introduction to the methodological design 26

1.4.1 Project stages flowchart 27

1.4.2 Ethical approval 29

Chapter 2 – Systematic review of the literature **30**

2.1 Introduction 30

2.1.1 Rationale behind the systematic review 30
2.1.2 Systematic searching for evidence in medical education 30
2.2 Objectives 32
2.3 Methods 32
2.3.1 Development of search terms 32
2.3.2 Inclusion / exclusion criteria 34
2.3.3 Database searches 35
2.3.4 Grey literature searches 37
2.3.5 Search dates 37
2.4 Results 38
2.4.1 Search results by stage 38
2.4.2 Summary of studies that met inclusion criteria 41
2.4.3 Summary of studies found via citation and reference 45
2.4.4 Summary of studies which added context 46
2.4.5 Summary of studies found in August 2012 update 48
2.4.6 Summary of studies from the grey literature 49
2.5 Conclusions 51
2.6 Discussion 52

Chapter 3 – Developing the intervention and the research design 55
3.1 Introduction 55
3.1.1 Background to London Speciality School of Paediatrics 55
3.1.2 Background to trainee representation in paediatrics 56
3.1.3 Leadership learning in paediatrics and the rationale for work-based learning .. 56
3.2 Objectives 58
3.3 Methods 58
3.3.1 Outline of the leadership development intervention 59
3.3.2 Key principles of the leadership development intervention 60
3.3.3 Application to the trainee committee 61
3.3.4 Participation in the trainee committee 63

3.3.5 Project-based leadership learning 63
3.3.6 Workshop-based leadership learning 64
3.3.7 Designing a leadership 360 degree multi-source feedback tool 66
3.3.8 Designing the evaluation strategy 70
3.4 Results 78
3.4.1 Applications to the trainee committee 79
3.4.2 Participation in the trainee committee 80
3.4.3 Project work uptake 80
3.4.4 Workshop uptake 81
3.4.5 Leadership 360 degree multi-source feedback tool uptake 82
3.5 Conclusions 83
3.6 Discussion 83

Chapter 4 – Quantitative measurement of the starting point 88
4.1 Introduction 88
4.1.1 Finding an objective measure 88
4.1.2 Self-assessment 90
4.1.3 Vignettes 92
4.2 Objectives 93
4.3 Methods 93
4.3.1 Development of the self-assessment tool 94
4.3.2 Piloting 99
4.3.3 Implementation 100
4.3.4 Statistical analysis and rationale 101
4.4 Results 104
4.4.1 Response rates 104
4.4.2 Comparison of trainee cohort and comparator group 105
4.4.3 Using factor analysis to determine summary themes 107
4.4.4 Progression through training 108
4.4.5 Does the experience log correlate with self-assessment scores? 109

4.5 Conclusions 111
4.6 Discussion 112

Chapter 5 – Qualitative analysis of the starting point 117
5.1 Introduction 117
5.1.1 Introduction to focus groups 117
5.2 Objectives 118
5.3 Methods 118
5.3.1 Data collection 119
5.3.2 Data analysis 120
5.3.3 Familiarisation & emergence of themes 120
5.3.4 Identifying a thematic framework 121
5.3.5 Indexing 121
5.3.6 Charting 121
5.3.7 Mapping & interpretation 121
5.4 Results 122
5.4.1 Leadership qualities in doctors as leaders 124
5.4.2 Leadership behaviours in doctors as leaders 126
5.4.3 Opportunities to support leadership development in the workplace 130
5.4.4 The broader aspects of leadership training and learning 136
5.5 Conclusions 138
5.6 Discussion 139

Chapter 6 – Quantitative measurement of the end point 144
6.1 Introduction 144
6.2 Objectives 144
6.3 Methods 144
6.3.1 Use and implementation of self-assessment tool 145
6.3.2 Statistical analysis and rationale 145
6.4 Results 148

6.4.1 Response rates 148
6.4.2 Comparison of pre- and post-intervention scores 148
6.5 Conclusions 151
6.6 Discussion 151

Chapter 7 – Qualitative analysis of the end point 156
7.1 Introduction 156
7.2 Objectives 156
7.3 Methods 156
7.3.1 Data collection 157
7.3.2 Data analysis 158
7.4 Results 158
7.4.1 Developing team-working, networking and collaborative skills 160
7.4.2 Establishing new approaches to personal development 163
7.4.3 Understanding the importance of the wider context 164
7.4.4 Multi-source feedback tools in supporting leadership development 166
7.4.5 Suggested improvements to the design of this programme 168
7.5 Conclusions 170
7.6 Discussion 171

Chapter 8 – Summary discussions 174
8.1 Summary of findings 174
8.2 Analysis of findings 177
8.2.1 Strengths of the research 178
8.2.2 Weaknesses of the research 180
8.3 Next steps 183
8.3.1 Next steps – longer term follow-up 184
8.3.2 Next steps – wider implications of this work 184
8.3.3 Next steps – implementation of learning from this project 186
8.4 Personal learning from this work 188

8.5 Future research 189

References 190

Appendices 203

List of Figures and Tables

Figure 1.1 Flowchart detailing the key steps within the overall study

Figure 2.1 Final search terms for systematic review

Figure 2.2 Summary of numbers of papers found at each stage of the main search of the databases

Figure 3.1 Outline of the 4 components of the leadership development intervention

Figure 3.2 Introductory page of 'Leadership 360' multi-source feedback tool

Figure 4.1 Miller's framework of clinical assessment

Figure 4.2 Domains of the Medical Leadership Competency Framework

Figure 4.3 First 3 sections of pre-intervention self-assessment questionnaire

Figure 8.1 Outline of the four components of 'Paired Learning'

Table 1.1 Domains and elements of the Medical Leadership Competency Framework

Table 3.1 Kirkpatrick's four levels of training evaluation

Table 3.2 Categorisation of successful applicants to the trainee committee by level of training, hospital where they were based and sub-group placement

Table 4.1 Domains and elements of the Medical Leadership Competency Framework (MLCF) and their relationship with project 'Dimensions'

Table 4.2 Number of self-assessment questionnaire respondents by training grade

Table 4.3 Comparison of pre-intervention self-assessment scores between the comparator and trainee committee groups.

Table 4.4 Results of dimension reduction using factor analysis

Table 4.5 Pearson's product-moment correlation coefficient analysis of correlation between self-assessment questionnaire scores and number of years in specialty training

Table 4.6 Multivariate ANOVA detailing the relationship between Likert self-assessment scores and documented previous experiences

Table 4.7 Experiences that had a significant influence on increasing self-assessment scores for each of the 11 dimensions

Table 5.1 Allocation of pre-intervention focus group participants with illustration of the number of each participant's quotes used in the thematic analysis

Table 6.1 Comparison of pre-intervention and post-intervention self-assessment mean scores in the trainee committee group.

Table 6.2 Comparison of pre-intervention and post-intervention self-assessment mean scores in the trainee committee group using paired t-tests

Table 7.1 Allocation of post-intervention focus group participants with illustration of the number of each participant's quotes used in the thematic analysis

Chapter 1 – Background

Chapter Overview

This chapter explains the context and background to the thesis. It begins by describing the theoretical perspectives of leadership on which this work is built. The chapter then describes the background to clinical and medical leadership within the NHS, and introduces the Medical Leadership Competency Framework (MLCF). It then details the aims and objectives of the thesis and maps out each of the exploratory methods that are used within the thesis on a flowchart.

1.1 Background to wider leadership literature and theories

This section provides a theoretical basis for the thesis, through a critique of some of the key themes and arguments from the wider leadership literature.

1.1.1 Introduction to the evolution of theories of leadership

The study of leadership has existed for many years and has been consistently described and documented since the 1920s, when it began to be an area of academic study. It can be argued that the emergent leadership theories over the last 80-90 years have been influenced by the society in which they are set (1), which makes the case for the importance of new leadership development initiatives being relevant to the society in which they are being introduced. Initially, the main emphasis of academic leadership theorists was on the individual, with focus on describing the traits and personalities of great leaders. Their approach subsequently moved to explore the behaviours of leaders, underpinned by the generation of empirical evidence through observation of what leaders actually did. In recent years there has been a more explicit emphasis on a leadership approach that works within a society where social hierarchies have been broken down. This evolution through some of the more widely recognised leadership theories, and how this has provided a theoretical framework for this thesis, is described below.

1.1.2 Trait theory

Much of the early leadership literature focused on trying to identify and describe the characteristics of individuals, who by their status, power or reputation were considered to be leaders. This approach, known as the trait theory, assumed that these individuals possessed certain traits or characteristics that made them ‘natural’ leaders. If these traits, which were also regarded by many academics as innate, could be identified in people, it could be argues that it

would enable them to be appointed to the appropriate leadership roles. Despite extensive academic pursuit of the key characteristics only a few factors – intelligence, self-confidence, drive, integrity and achievement – have emerged with any stability across different research contexts (2)(3), and so many theories of leadership began to more towards a more behavioural framework (4). Some of the difficulties with using trait theories to shape an approach to leadership development are that, firstly, there needs to be consensus about which characteristics should be considered the most important, and, secondly, the characteristics need to be stable across different situations (1). Unfortunately, it can be argued that neither of these really hold true (5)(4). Spurgeon, Clark and Ham suggest that the trait approach might have some resonance with the medical profession because of the focus placed on key personal characteristics in selection into medicine, and in recruitment to specialty training programmes. However, they point out that while doctors may employ personal qualities in the patient-doctor context, the same traits may not necessarily have the same weight or influence in the dynamic group context of leadership (2).

Despite these difficulties with context and situation, interest in trait theory remains. Alimo-Metcalfe and Alban-Metcalfe (6) have written widely around a new set of descriptors that they have studied through questionnaire-based research – approachable, inspirational, genuine concern, collaborative. They argue these should be more accommodating of a more modern context. However, the issues around weighting and stability, described in the paragraph above, remain. There are also difficulties in how one measures traits (7). With these concerns around instability across different situations, trait theory has not been actively used to shape the medical leadership intervention in this work. However, the nature of the focus groups used in the research may allow for discussions around leadership traits to emerge within this exploratory research.

1.1.3 Leadership styles and behaviours

In response to the criticisms of the failure of trait theories to recognise the impact of different contexts or situations in which leadership occurs (8), leadership researchers began to describe new leadership models that focused on leadership behaviours, or styles. Two broad sets of behaviours emerged; task behaviours where there is priority on getting different aspects of the job done, and relationship behaviours where the emphasis is on the way people work together (9)(10)(11). The aspiration of this research was to link a specific style to a measurable outcome. However, as with much of the trait theory work, it can be hard for a clear relationship to be established (12). Whether or not task and relationship styles can be found in one person, can play out simultaneously, can realistically adapt to the changing context and

can be the useful target of leadership development interventions are areas of debate (13)(14). There is criticism that, as with trait theories, leadership-style theories fail to consider the leadership setting (15)(16), and although they focus on behaviour, they do not directly address the values that are relevant and effective in delivering outcomes. Derue et al (17) performed a meta-analysis of 79 previously published studies in an attempt to explore the evidence behind how traits and behaviours interact with each other, and to determine whether one approach is better than the other. They conclude that overall leadership behaviours are a better predictor of leadership effectiveness that traits, and that the most effective leaders are those who plan work, support and enable their followers, and facilitate change.

Goffee and Jones have viewed the trait and behavioural theories through a different lens, arguing that it is not the leadership style that makes a great leader, but the underlying authentic personal qualities that make the style effective (18). This approach to the trait and behavioural theories of leadership informs the first domain of the Medical Leadership Competency Framework (MLCF) (19), which is described in the section below, and has an important impact on the design of the leadership development intervention in this research.

1.1.4 Transactional and transformational leadership

The realisation that the external environments in which leadership challenges arise are variable and unstable led to the emergence of a sizeable literature exploring the concepts around transactional versus transformational leadership (20)(8)(21)(7). Transactional leadership can be described as a series of exchanges between a leader and his or her followers, normally based on a hierarchical organisational structure. By contrast transformational leaders release potential through the empowerment and development of their followers, although some commentators argue that the two can be viewed as a continuum (21). This relationship between leaders and followers is crucial to these theoretical approaches (22). In transformational leadership organisational vision and values are clearly stated and the work of each individual within the team is aligned to the achievement of longer-term goals (23). Transformational leadership has been a widely-adopted model, and has been incorporated into many public sector leadership frameworks across the world. It has also strongly influenced the shape of the Medical Leadership Competency Framework (MLCF) (19), which in itself has provided an important framework for this thesis. It can be argued that the main disadvantage of this is the focus is on the individual leader, rather than the organisation as a whole (7)(24).

1.1.5 Healthcare organisations as professional bureaucracies

Another influential leadership theory comes from the work of the organisational theorist Henry Mintzberg, who characterised healthcare organisations as 'professional bureaucracies' rather than 'machine bureaucracies' (25). In professional bureaucracies, of which healthcare organisations would be a key example, the front-line staff members, he argues, have a large measure of control of the nature and content of their work. It is their education, training and specialist expertise that gives them that influence. This means that compared to a machine bureaucracy, such as a government department, directives passed down from the top of a 'chain of command' often have relatively little impact, and may at times be ignored or resisted by frontline staff. Mintzberg highlights that, furthermore, professional bureaucracies have an inverted power structure where staff members at the bottom of the organisation often have greater influence over the day-to-day decision making than other staff in more senior positions. From this emerges the important concept of 'followership' (26). The influence of frontline staff within a professional bureaucracy provides both challenges and opportunities for leaders in these organisations. While organisations can benefit from these 'bottom-up' influences, Badrick and Preston's (27) examination of the implementation of quality management systems in a number of healthcare settings showed that it was the professional bureaucracies that struggled the most to implement a consistent approach to quality. It can be argued that this challenges our current approach of regulation and performance management within the health system in the UK.

Spurgeon, Clark and Ham argue that three implications for leadership follow on from Mintzberg's theories (2). Firstly, in professional bureaucracies, professionals and the networks that they form play an important leadership and co-ordination role. Peer support, peer review and, on occasions, peer pressure are all important horizontal processes that can hold great influence. Secondly, they argue that professional bureaucracies are characterised by distributed leadership where large numbers of clinical leaders from different backgrounds and at different levels work together. Clinical microsystems are important systems where this distributed leadership can sit (28). Thirdly, they highlight the importance of collective leadership in healthcare organisations, where leadership comes from teams rather than necessarily just charismatic individuals. These implications are highly relevant in terms of their support for a work-based team-orientated approach to leadership development, where groups of clinical leaders from different levels and backgrounds come together to lead improvements. The ideas around peer-support and distributed team-based leadership are important influences on the design of the leadership intervention in this research.

1.1.6 Shared Leadership

One of the criticisms of the transformational leadership approach described in sub-section 1.1.4 above is that focusing on one individual is somewhat limited (7)(24). Within healthcare, increasingly complex problems and sets of tasks facing organisations have led to the emergence of the delivery of care by multiprofessional teams. Leadership tasks are shared across teams, networks and organisational boundaries. Shared leadership is an approach that can support and enable this (29), and is crucially much more focused on *leadership*, rather than on a specific leader. Leadership is distributed amongst a group of individuals rather than being held centrally by a single person who acts as the leader (30). The knowledge, capability, experience and capacity of each individual member of a team are highly valued and together are used to share, or distribute, the different leadership tasks across a team. There is a growing body of evidence that suggests that a shared leadership approach can deliver success across a range of different sectors (31)(32). Specifically, there is evidence that shared leadership can increase risk taking, innovation and commitment, which should improve the performance of a team, and result in improved care for patients (31)(33). Ulhoi and Muller (34) analysed 271 papers relating to shared, collective and distributed leadership. They concluded that if shared leadership is to be seen as a genuinely collective phenomenon then the leader-follower dichotomy of leadership needs to be replaced. They argue that with the right support and development individuals are capable of playing both the leader and follower roles.

The Medical Leadership Competency Framework (MLCF) (19), which is described below in sub-section 1.5.4, and is the leadership framework used in this thesis, is built on these theoretical principles of professional bureaucracies and shared leadership.

1.2 Background to clinical leadership and medical leadership

This section builds on the leadership theories discussed above, to focus in on clinical leadership and then more specifically medical leadership. It also introduces the Medical Leadership Competency Framework (MLCF) (19), and describes the lead researcher's (RK) involvement in the national work to embed the framework within postgraduate medical curricula across the UK. This experience, as explained in Section 1.2.4, had a significant impact on the development of this thesis.

1.2.1 What is clinical leadership?

'Clinical leadership' is a concept that has emerged from these broader leadership theories (23). It is a phrase that has become increasingly used within healthcare in the last 5-10 years, yet for

many of the 1.4 million employees of the NHS (35) there is the potential for uncertainty as to what this actually means for them and the patients they look after. Mountford and Webb's illustration of a number of global examples of where clinical leadership has positively impacted on clinical outcomes provides a pragmatic, case-based definition (36):

> 'Clinical leadership is putting clinicians at the heart of shaping and running clinical services, so as to deliver excellent outcomes for patients and populations, not as a one-off task or project, but as a core part of clinicians' professional identity (37).'

There is growing case-based recognition that strong clinical leadership, shared in a distributed way across clinical teams, is an effective way to manage and improve services in order to deliver the very best possible care to patients (2)(38)(39). Some of the strongest examples of where clinical leadership is embedded throughout organisations come from the Mayo Clinic (40) and Kaiser Permanente (41), in the United States. There are also UK examples emerging where clinical leadership has been a key driver in improved clinical outcomes (39)(42). In each of these examples, the conclusions of the authors are that it is the partnerships between doctors and managers, the team-based approach where individuals can have roles that move between followers and leaders, and the focus on quality that have led to improved outcomes.

1.2.2 Clinical leadership in the NHS

Within the NHS, clinical leadership has emerged in recent years as an important policy driver to the organisation and delivery of safe and efficient healthcare. In the UK, Lord Darzi's NHS Next Stage Review in 2008 described clinical leadership as:

> 'Central to our expectations of the healthcare professionals of tomorrow (43)'

by which I think he is demanding that clinical leadership will be something that is expected of everyone working in the NHS. Prof Sir John Tooke, in 'Aspiring to Excellence', the inquiry into the Modernising Medical Careers programme that sought to transform postgraduate medical training in the UK, recognised the need to bring leadership into education and training, although it could be argued that the word 'commander' gives an emphasis removed from the theories around shared and collective leadership described above:

> 'The doctor's frequent role as head of the healthcare team and commander of considerable clinical resource requires that greater attention is paid to management and leadership skills regardless of specialism (44).'

Ham and Dickinson conducted two reviews in 2007 to explore what could be learned from the international experience and research evidence in support of engaging doctors in clinical leadership (45). The first was a survey where experts in different countries were asked to write

a paper on their experiences, and the second was a more systematic review of the role of medical engagement in the leadership literature. In their paper they argue for a concerted focus on developing the training and support for doctors in taking on leadership roles. Although much of this work emerged from professional experience rather than empirical evidence, the breadth of healthcare systems examined gives their analysis strength.

In their 2012 guidance, 'Leadership and management for all doctors', the GMC recognises that while doctors have historically been trained to care for the patient who is in front of them, there has been less emphasis on developing a knowledge of the wider system in which they work (46). This guidance, which highlights the role of the doctor in leading system-wide approaches (ie ones that reach beyond individual organisations, specialties or patient pathways) to delivering safe and effective care, was established through the formation of a working group and then a process of formal consultation. This is important as it provides a base from which undergraduate and postgraduate curricula, as well as guidance around appraisal and revalidation, can be influenced.

In his New England Journal of Medicine paper on clinical leadership (47) Richard Bohmer highlights survey-based work that suggests clinicians want a greater leadership role but feel unprepared or disempowered (48). He argues that healthcare leaders can encourage and support clinical leadership by framing their organizational purpose as creating value, and through giving local leaders the authority to make microsystem changes.

There are other researchers who argue that developing experience and capabilities in these areas is important for all clinicians, as we all have a responsibility to constantly seek to improve the system within which we work. In his paper based on interviews and site visits to over 20 clinical microsystems in the United States Paul Batalden developed the argument as to how 'quality improvement' can transform healthcare (49), concluding that:

> 'everyone in healthcare really has two jobs when they come to work every day: to do their work and to improve it.'

This is important as it sets out the importance of creating a culture of continuous improvement amongst all staff within different healthcare organisations.

In 2013 Dickinson et al published a mixed method study to examine the current evidence for the impact of clinical leadership in England (50). This involved a questionnaire survey of NHS trusts in England, case studies of nine NHS trusts that responded to the survey; and the use of

the Medical Engagement Scale in these case studies to establish the extent to which doctors feel engaged in the work of their organisations. They found many barriers to involving doctors effectively in leadership roles, and concluded that in most organisations a significant change in approach is needed to overcome these barriers. Their ideas included increasing the time commitment of medical leaders, positioning more doctors in formal leadership roles and developing a culture of engagement that was found in the most leading organisations. This is important as it begins to provide an evidence-base to support policy makers and leadership development leads to develop programmes with the best possible approach.

A challenge for the NHS, and for other healthcare systems across the world, is how to train, enable and support their workforce to develop as clinical leaders and to be able to take on both of these tasks. In the UK, the establishment of the Faculty of Medical Leadership and Management (FMLM) in 2011 has been one national-level approach to tackling this challenge. FMLM is a professional membership organisation which aims to promote the advancement of medical leadership, management and quality improvement at all stages of the medical career for the benefit of patients. Looking beyond doctors and medical leadership, the NHS Leadership Academy was established in England in 2012 with the principal purpose of developing outstanding leadership in health, with a continual focus on improving patients' experiences and their health outcomes. Although very young, both of these organisations are working hard to evidence the impact of their leadership development work. To date this has been through the publication of case studies (51) and outputs such as the consultation for and publication of professional standards (52), but as their influence and networks widen their ability to influence policy, education and training curricula and service standards is likely to strengthen.

1.2.3 Clinical leadership or medical leadership?

Most of the concepts discussed within the previous section have related to clinical leadership, which has applicability to all clinicians or health professionals working together to improve care for patients. The subjects involved in this thesis are postgraduate medical trainees in paediatrics, ie doctors, so there is an argument that the term ‘medical leadership’ should be used. Indeed, Spurgeon, Clark and Ham, in their highly informative text on medical leadership and engagement (2), argue that too often commentators use the term ‘clinical leadership’, when clearly meaning ‘medical leadership’. They point out that doctors have been involved in the running of health services, locally, nationally and internationally since the very first pioneers established and ran health services many centuries ago. They also highlight the emergence of evidence of the relationship between the extent to which doctors are engaged in

the planning, shaping and leading of services and the wider performance of their organisation (53). They do acknowledge, though, that the majority of policy statements, papers and directives prefer to use the term clinical leadership, despite their focus on doctors. Perhaps it is the need to contextualise this leadership development within a modern health service, where the delivery of care is a multiprofessional team activity, that leads to a preference to use 'clinical leadership'. In this thesis, which is a study of leadership development in doctors, using the term 'clinical leadership' is the approach that has been taken, although with the introduction of the Medical Leadership Competency Framework (MLCF) in the next section it must be acknowledged that the terminology is on reflection unfortunate.

1.2.4 Medical Leadership Competency Framework (MLCF)

The Medical Leadership Competency Framework (MLCF) was developed through collaboration between the NHS Institute of Innovation and Improvement and the Academy of Medical Royal Colleges (19). It was based on a detailed review of medical leadership experience from across the world, and through a comparative analysis of medical curricula and leadership competency frameworks including the NHS Leadership Qualities Framework (LQF), which was introduced in 2006 to support the development of senior NHS leaders (54). The project team who designed the MLCF also consulted with members of the medical and wider NHS community and received input from a project steering group, a reference group and focus groups. Since it was originally published in May 2008, the MLCF has been tested in a range of medical education and service communities feedback has been received from many individuals and groups including patient and lay groups, equality and diversity experts, medical students and doctors, and from organisations implementing the framework. Many of the suggested improvements have been incorporated in to the second edition and third editions.

Shaped by the input of this wide range of stakeholders, the MLCF aimed to describe the 'competencies' that doctors need to develop as clinical leaders. My interpretation of the word competencies in this context are that these are the skills, approaches and behaviours that underpin clinical leadership. These competencies, which were arranged into 5 domains, each with 4 different elements, had a focus on developing personal qualities, working within teams, and becoming actively involved in the planning, delivery and transformation of clinical services.

MLCF Domains:	MLCF Elements:
Demonstrating Personal Qualities	Developing Self Awareness
	Managing Yourself
	Continuing Personal Development
	Acting with Integrity
Working with Others	Developing Networks
	Building & Maintaining Relationships
	Encouraging Contribution
	Working within Teams
Managing Services	Planning
	Managing People
	Managing Resources
	Managing Performance
Improving Services	Ensuring Patient Safety
	Critically Evaluating
	Facilitating Transformation
	Encouraging Improvement & Innovation
Setting Direction	Identifying the Contexts for Change
	Applying Knowledge and Evidence
	Making Decisions
	Evaluating Impact

Table 1.1 Domains and elements of the Medical Leadership Competency Framework (19)

Reflecting back it is clear that the development of personal qualities relates back to the discussions around leadership styles in sub-section 1.1.3 (18), and that elements within the working with teams, managing and improving services domains are strongly underpinned by the wider leadership research evidence discussed in the sub-sections on transformational leadership, healthcare organisations and professional bureaucracies and shared leadership. These theories have influenced the shape and tone of the MLCF, and the different competencies detailed within it. It could be argued that the MLCF provides a language for doctors in training, and those who train them, to access some of the learning and application of these leadership theories.

During the period when I was designing and planning this research project I was asked to join the 'Enhancing Engagement in Medical Leadership' project team who were responsible for the development of the MLCF. At this time the 1st edition of the MLCF had been published so much of the focus was on implementation, where there were two main challenges to address. The first of these was to work to influence the key bodies involved in setting standards,

curricula and managing the quality assurance processes (ie the General Medical Council, the Postgraduate Medical Education Training Board and the Medical Royal Colleges) to increase the presence of medical leadership and the MLCF within their work. The second, which aligned directly to my work in this research project, was focused on enthusing and enabling trainees and trainers to engage with the MLCF and to use it as a framework and stimulus to make opportunities for work-based leadership learning.

My experiences while undertaking this work were of real significance to this thesis. During the time I was working with the team the 2nd and 3rd editions of the MLCF were published (19). We also succeeded in incorporating it into the GMC's guidance for undergraduate doctors Tomorrow's Doctors (55), developed guidance for medical schools, and developed the Medical Leadership Curriculum (56) which was incorporated into the 58 specialty curricula of the Medical Royal Colleges and Faculties. Above all this work gave me a deep understanding of what clinical leadership really means in practice. It allowed me to work through how written competencies could be learned and assessed within the workplace, and to work with clinical colleagues to explore which activities could best provide leadership learning. I also found the space to reflect back on my understanding of the wider leadership literature and to work through the possibilities around developing a work-based leadership development intervention within this study.

In addition, the MLCF gave important content validity to the self-assessment questionnaire and scale that were developed for this project, and provided a strong framework for much of the content of the leadership development intervention. Chapter 4 describes the opportunities and limitations of this in more detail.

1.2.5 Leadership learning within paediatrics

Paediatricians are the clinical group considered in this thesis, with the question around whether they can learn clinical leadership skills. It therefore seems important to briefly describe the context of their work and some of the current issues within the specialty. Paediatrics has a number of important challenges that it is currently facing. These include increasing child protection work, large-scale public health initiatives around immunisation and obesity, and the development of adolescent-focused services to support the transition of the care for children and young people from paediatric services and into adult services. With moves towards a model of Consultant-delivered care there are complex workforce planning issues and concerns around potential recruitment difficulties in the future (57)(58). While all of these issues generate challenge and anxiety for paediatricians, they also provide opportunities for them to

develop a greater understanding of some of the wider, systemic issues within healthcare, and to develop as clinical leaders. These challenges may be different for other specialties although the concept of a doctor learning to move on from just thinking about the single patient within a consultation to also considering a more systemic approach applies in all cases.

It has been interesting to also reflect back on some of the leadership theories discussed earlier in the chapter to consider their influence on the way in which paediatric services are led and run. Paediatrics has a long and proud history of multi-professional working (59)(60), with many departments running their services through the collaborative work of multi-disciplinary teams (MDTs). This approach to running paediatric services strongly relates to the theories around professional bureaucracies (25) and shared leadership (30). The nature of this work provides opportunities to develop self-awareness, to encourage and acknowledge the contribution of others and to learn team working skills, all of which are key leadership competencies highlighted within the MLCF.

1.2.6 Definitions of leadership

Leadership has been described as a phenomenon that almost everyone has an opinion on, but definitions of leadership are often strongly contested and tend to be context-bound (61). As Bernard Bass observed, there are almost as many different definitions of leadership as there are persons who have attempted to define it (8). The interpretation that Peter Spurgeon and I used in our medical leadership book, aimed at tutors and trainees, was that:

> ‘contemporary leadership seems to be a matter of aligning people towards common goals and empowering them to take the actions needed to reach them (12).’

Our rationale for the book was to use a simple, workable definition that would be accessible to colleagues working within clinical settings, and this argument also applies to this research. It is important that the paediatrician participants in this study have a definition of leadership that resonates with them as they look to gain leadership learning.

The other confounder to briefly consider is the debate around the differences between leadership and management. One school of thought has emphasised that leadership and management are two completely different functions. Zaleznik (62), for example, has suggested that leaders develop visions and drive changes while managers monitor progress and solve problems. In their introduction to clinical leadership, Swanwick and McKimm highlight some of the differences between leadership and management, before pointing out one of the pitfalls of taking an approach that separates them out:

> '....some unique features of both activities, namely leadership as being about setting direction, influencing others and managing change: with management concerned with the marshalling and organisation of resources and maintaining stability. This approach tends to denigrate management as boring and unsatisfying: who would want to 'manage' when they can 'lead'? (23)'

In reality, most significant tasks in organisations require aspects of leadership and management. There are perhaps best thought of as processes which interact and support each other. They are both essential for effective organisations, but at different times one might be emphasised more than the other (63). For this reason this more integrated interpretation of leadership and management is one that has been adopted within this research.

1.3 Aims and objectives of the thesis

1.3.1 Overall aim of the thesis

The overall aim of this thesis was to explore whether paediatricians in training can develop leadership skills through participating in a specifically designed leadership development initiative.

1.3.2 Overall thesis objectives

The objectives of this thesis were to:

a) Perform a focused systematic review of the medical leadership literature, and to analyse key learning and themes from these leadership development initiatives.
b) Design and implement an innovative leadership development initiative to support the learning and development of a newly established postgraduate trainee committee, and design a research strategy to evaluate it.
c) Develop a self-assessment tool to measure the leadership capabilities of the trainee committee cohort prior to the implementation of the leadership development intervention, and to compare these results with a comparator group of trainees.
d) Explore participants' ideas around leadership to gain a deeper, qualitative understanding of the baseline leadership learning of the trainee committee cohort, prior to the implementation of the leadership development intervention.
e) Evaluate quantitatively the effectiveness of the leadership development intervention experienced by the trainee committee cohort.
f) Analyse qualitatively the effectiveness of the leadership development intervention experienced by the trainee committee cohort.
g) Summarise the findings from the project, highlighting areas of personal learning, potential wider implications of the project and suggestions for future work.

These seven objectives are approached in turn by each of the chapters 2-8, an outline of which is presented below.

1.3.3 Orientation of chapters

This sub-section aims to orientate the reader around the subsequent seven chapters, and to illustrate how they come together as a thesis:

Chapter 2 – ‘Systematic Review of the Literature’ (page 30) details the systematic review of the medical leadership literature. The purpose of this review was to explore the empirical evidence around leadership development in medicine. It describes the search methodologies used, discusses the results of these searches, and explores the themes and learning that can be drawn from the review.

Chapter 3 – ‘Developing the Intervention and the Research Design’ (page 55) focuses on the development of the leadership intervention and outlines the design of a mixed methods research methodology used to evaluate it. The chapter describes the vision of a trainee committee within the London Speciality School of Paediatrics and Child Health, and discusses how their work, and the experiences gained, underpinned a work-based leadership development programme. The chapter details the selection of the trainee committee cohort, and discusses the design, refinement and implementation of the leadership intervention itself. It also describes the trial of a leadership multi-source feedback tool which aimed to support the learning and development of the participants who wanted to use it. The chapter also provides an overview of the mixed methods research methodology used in this thesis, and describes the rationale behind it.

Chapter 4 – ‘Quantitative Measurement of the Starting Point’ (page 88) describes the quantitative methodologies used to measure the pre-intervention baseline (pre-test) or ‘starting point’. It introduces the ‘comparator’ group of trainee paediatricians against which the trainee committee cohort could be compared, and describes the development of a self-assessment tool that was used for this measurement. The chapter explores the use of vignettes within the self-assessment tool, the attainment of content validity, the development of an ability scale and summarises some important issues around self-assessment. The chapter details the results of these ‘experiments’ and discusses the conclusions that can be drawn from them.

Chapter 5 – ‘Qualitative Analysis of the Starting Point’ (page 117) builds on the quantitative assessment of the previous chapter. It introduces the principles of framework analysis as a qualitative research methodology. It continues by describing the focus groups that were used to explore participants’ ideas around leadership, and to gain a deeper, qualitative understanding of the baseline leadership learning of the cohort of participants, prior to the implementation of the intervention. The addition of this qualitative research ‘experiment’ gives the overall project a strong mixed-methods approach. The chapter details the thematic analysis of this work and explores the conclusions that can be drawn from it.

Chapter 6 – 'Measurement of the End Point' (page 144) examines the quantitative methodologies that were used to measure the post-intervention (post-test) 'end point' of the study and hence evaluate the impact of the leadership development intervention. The chapter details the comparative results of the pre-intervention and post-intervention self-assessment scores, describes the rationale behind the different statistical analyses used in these comparisons and explores the conclusions that can be drawn from them.

Chapter 7 – 'Qualitative Analysis of the End Point' (page 156) expands on the quantitative methodologies that have been used to measure the post-intervention end point, by describing the focus groups that were run at the end of the first year of the programme. These focus groups were designed to capture a deeper understanding of the more qualitative aspects of the leadership learning gained from being a participant on the programme. The chapter also describes other themes and learning that emerged and includes a qualitative analysis of the use of the leadership multi-source feedback tool described in chapter 3. The chapter details the formal thematic analysis of this qualitative work and explores the conclusions that can be drawn from it.

Chapter 8 – 'Summary Discussions' (page 174) focuses on summarising and analysing the findings from the mixed method approaches described in the previous six chapters, and then contextualising them within the wider healthcare leadership literature. It examines the impact of the intervention, the strengths and weaknesses of the research, and synthesises the findings whilst exploring the wider implications of the work. The chapter contains some reflections on my personal learning from this project, before moving on to consider possibilities for longer term follow-up and the applicability of the findings to other groups. The chapter concludes by summarising the key conclusions and discussing how the learning from this project might be implemented into other leadership development initiatives.

1.4 Introduction to the methodological design

This project aimed to explore whether trainee paediatricians could develop clinical leadership capabilities through their participation in a specifically designed leadership development programme. In order to explore this question in detail a mixed methods quasi-experimental design was used. This over-arching approach, which is recognised as being valuable within health services management (64) and medical education (65) is succinctly described by Schifferdecker and Reed (66):

> Mixed methods research involves the collection, analysis and integration of both qualitative and quantitative data in a single study. The benefits of a mixed methods

approach are particularly evident when studying new questions or complex initiatives and interactions, which is often the case in medical education research.

Key to this description is the acknowledgement that mixed methods research is more than just collecting and analysing data about a particular subject from more than one angle. The important step of integration of the data (67) allows for more focused conclusions to be me made. This is discussed in further detail later on in the thesis.

The methods used in this thesis were theoretically and empirically informed, and together contribute to the mixed methods approach that was chosen:

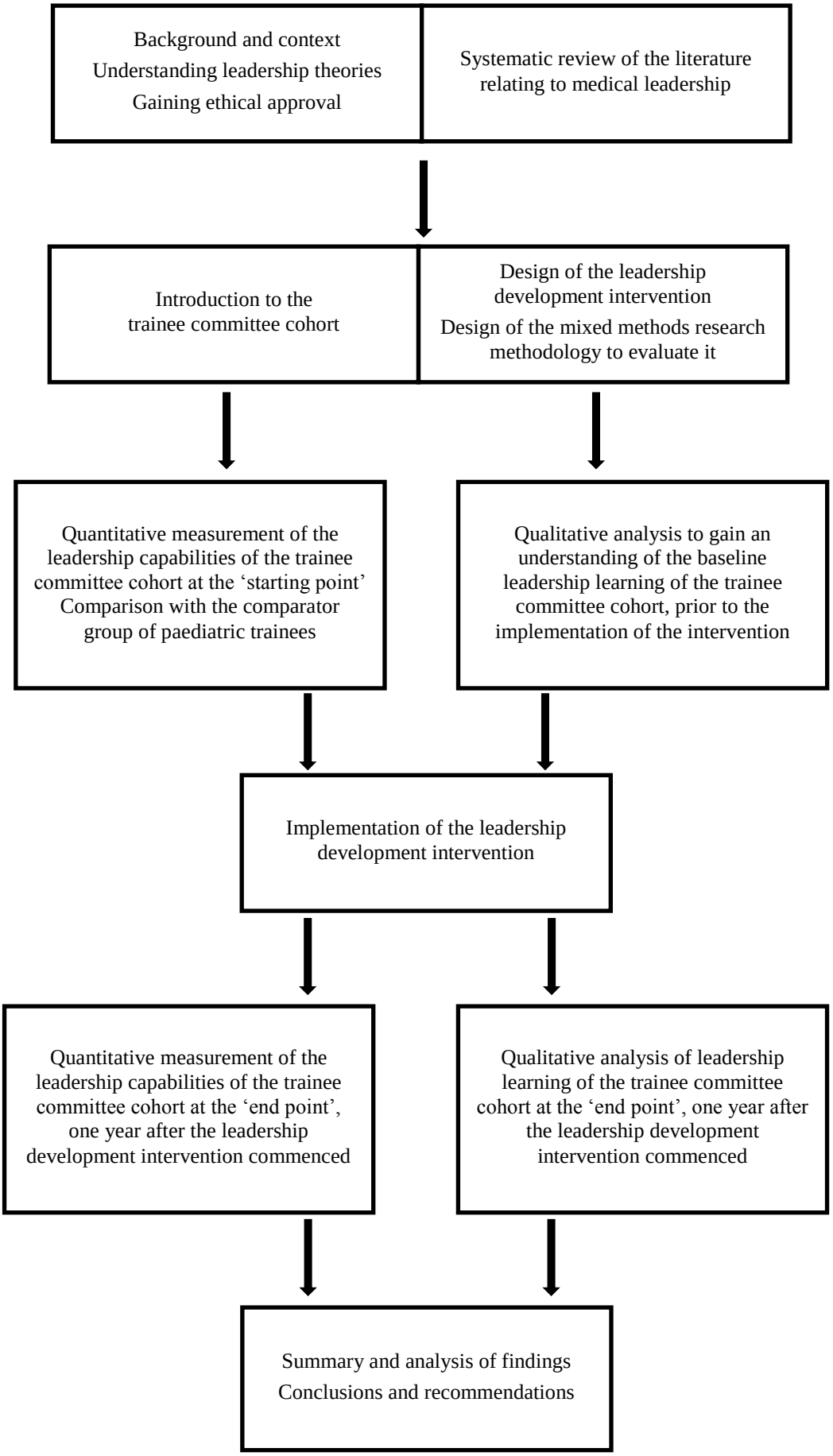
1. Systematic review of the medical leadership literature – this is described in detail in chapter 2.
2. Quantitative measurements using the self-assessment leadership capability surveys – these are described in detail in chapters 4 and 6.
3. Qualitative analyses of the trainee committee cohort's perceptions prior to, and after, the implementation of the leadership development initiative – these are described in chapters 5 and 7.

A more detailed critique of the design of and rationale for the mixed methods approach taken in this research can be found in sub-section 3.3.8 and section 3.6 within chapter 3. The detail of each of the steps involved in this overarching approach is described within chapters 4,5,6 and 7.

1.4.1 Project stages flowchart

Figure 1.1, on the following page, is a flowchart that highlights the key stages of the study. Each of these stages is represented by a box, and each is referenced to the specific chapter where the detail of that stage is described. Where boxes are adjacent to each other, the stages were undertaken at the same time. The specific detail of this sequencing is important and is discussed in chapters 4,6 and 8.

Figure 1.1 Flowchart detailing the key steps within the overall study:



1.4.2 Ethical approval

Whilst the risks of this educational research work were self-assessed as being small, and there was no direct involvement with patients, it was determined that ethical approval should be sought for this thesis for the following reasons:

a) The study involved NHS staff and there may be potential ethical issues that need considering prior to their participation
b) Participating in the ethics application process would provide learning for the lead researcher (RK) around the processes and rigor required when seeking approval
c) The study design would be strengthened by the peer-review of the ethics panel
d) Leading medical education journals suggest that medical education research should have formal ethical approval

A favourable ethical opinion on the research was granted on 9th March 2009 and is reported in Appendix 1.1.

Chapter 2 – Systematic review of the literature

Chapter Overview

This chapter builds on the outline exploration of the theoretical frameworks from the wider leadership literature in Chapter 1 to detail the systematic review of the medical leadership literature. It describes the objectives of the systematic review, before going on to give a detailed description of the search methodologies used. These include searches of medical, education and social sciences databases as well as an exploration of the grey literature. The results of these searches for empirical evidence are analysed and discussed in detail. The chapter concludes by exploring the themes and learning that can be drawn from the review, and discusses how these findings have influenced this work.

2.1 Introduction

This chapter builds on the theoretical discussions and introduction to clinical leadership from chapter 1 to describe the systematic review of the literature looking for empirical evidence of leadership development interventions in medicine.

2.1.1 Rationale behind the systematic review

As the aims and objectives of this thesis developed, the next stage was to re-visit the literature more formally through undertaking a full systematic review of the subject area. This review would systematically search across both the published and grey literature in order to find evidence of the impact of other medical leadership development initiatives. The findings from this review could then be used to shape the design, implementation and evaluation strategies of this project.

2.1.2 Systematic searching for evidence in medical education

With the emergence of web-based databases over the last two decades, the exponential increase in rapidly accessible data provides any researcher with a wealth of information within which to search for relevant evidence. The complexity arises from the sheer breadth and depth of this information; this can be very daunting when considering a search for evidence. Within the medical arena, bibliographic databases, which index publications with descriptors (often called subject headings), keywords and other identifiers, are the medium where the majority of evidence is likely to be found. However, in the field of medical education, the situation is more

complex as, although there are many databases dedicated to 'medicine' and others to 'education', there are very few comprehensive sources focused on 'medical education'. Whilst there are databases where one can search using keywords, these are reliant upon the accurate submission and indexing of keywords by authors and editors and this is often not the case (73). For this reason, the approach to a systematic review of a medical education focused research question has to be somewhat different to the exploration of evidence behind a clinical question. A helpful example of this type of approach can be found within a systematic review of factors contributing to nursing leadership, that was published in the Journal of Health Services Research and Policy in 2008 (74). This review includes an appendix which the authors used as a checklist to note down and score whether key points regarding the design, sample, measurement and statistical analysis were applicable to each retrieved paper. This enabled papers to be 'scored' as being of low, medium or high potential using an 'Overall Study Validity Rating'. A second appendix from this review was a 'Design Quality Assessment Tool' focused on assessing the quality of pre and post intervention studies. This tool generated an overall validity rating by asking key questions to assess the quality of the study in question in the following six areas:

a) Sampling
b) Design of the study
c) Control of confounders
d) Data collection and outcome measurement
e) Statistical analysis and conclusions
f) Drop-outs

The headings from this tool provided a helpful conceptual framework for this research, aiding the process of assessing whether papers found in the searches should be included.

As well as involving a wider set of databases, there is also a greater emphasis on the 'grey literature' within medical education and leadership reviews. Grey literature has been simply defined as "material which is not issued through normal commercial publication channels" (75) and can consist of a broad range of data. Examples would include reports, committee or conference proceedings, theses, PowerPoint presentations, newsletters, best practice documents and working group documents. Grey literature is often fairly succinct, which means it can be relatively easy to extract the relevant evidence from it. At the same time it may have been very heavily researched – this is often the case with government or technical reports where the grey literature version can have significantly more information in it compared to what is made available when the evidence is formally published (73). The grey literature was

often historically very difficult to locate, but the open nature of the internet had transformed researchers' access to these sources of evidence – the difficulty, with many fewer of these having had any peer-review, is knowing the provenance of the information (76). Sub-section 2.3.4 below details findings from the grey literature searches that were undertaken in this project.

2.2 Objectives

The key objective of this chapter is to describe the systematic review of the medical leadership literature that was undertaken in the early stages of this project, and to analyse the key learning and themes that emerged around leadership development initiatives within medicine. It is these themes that were used to shape the direction of this project and to develop the overall project objectives.

2.3 Methods

This section details the different steps used in the methodology behind the systematic review. In order to provide clarity of the processes undertaken, each stage is described under an individual sub-heading. Prior to commencing the systematic review time was spent carefully defining the research question, and after a number of iterations: "Do leadership development interventions improve leadership skills in doctors" was settled upon as the broad overarching question.

2.3.1 Development of search terms

Having defined the question, the next stage was to develop appropriate search terms that could be used in a wide range of databases to explore the research question. Guidance for this was taken from the Cochrane Handbook (77). The initial scoping considered possible search terms under the broader themes of medical leadership development, intervention, training or experience. This process was led by the lead researcher (RK) and supported by two junior researchers (RH and AR). In order to explore a breadth or options the research question was written in the centre of a large piece of paper and then each of the three researchers independently noted down essential subjects they judged to be related to the question. This was followed by noting down associated search terms. The three researchers then came together to compare outputs and to decide on a consensus version, which was signed off by the lead researcher RK.

Looking at the processes involved in more specific detail, the development of the search terms began by considering the participants. The scope of the question was around doctors, but it was decided to broaden the search out to medical students, in case there were studies or publications from medical schools with evidence of direct relevance to the postgraduate arena. Exploration of the literature revealed that there were some differences between search results when looking under 'doctor' compared with 'physician' so both were included as an 'either / or'. The word "consultant" was also considered, but with the breadth of the databases used, the wider connotation of the word led to a huge number of irrelevant 'hits' from journals like the Harvard Business Review. Other healthcare professionals were excluded from this search, as it was anticipated that their response to leadership development interventions may be significantly different to that of doctors. Therefore, for each of the searches, the initial set of search terms was for "Doctor" OR "Physician" OR "Medical Student". This in effect restricted the search to the medical leadership literature. The rationale for this decision was to ensure that any empirical evidence from the review could be relatively easily utilised within the leadership development intervention for the paediatric trainees. It has to be acknowledged that there may be evidence from other non-medical leadership development programmes that will have been missed by this approach, and that could be an important methodological weakness.

The second search term, which was focused on the intervention, and linked to the participant terms by the Boolean command AND, was much more difficult to finalise. After considerable experimentation with the organisation of essential subjects and associated terms by the three researchers described above, as well as discussions with the project supervisor (JD), four related search terms were settled on. Although there would be considerable overlap between all four, this diversity offset some of the risks of missing key papers when searching in databases where keywords and subject headings were not necessarily as tightly indexed as would be ideal. The four search terms were "Leadership training", "Leadership development", "Leadership intervention", and "Leadership experience". These two sets of search terms, each with their OR Boolean term, were combined together with an AND; ie, for a paper to be highlighted as a paper for further exploration it needed to be covered in both searches.

The summary of the final search strategy employed is given below in Figure 2.1

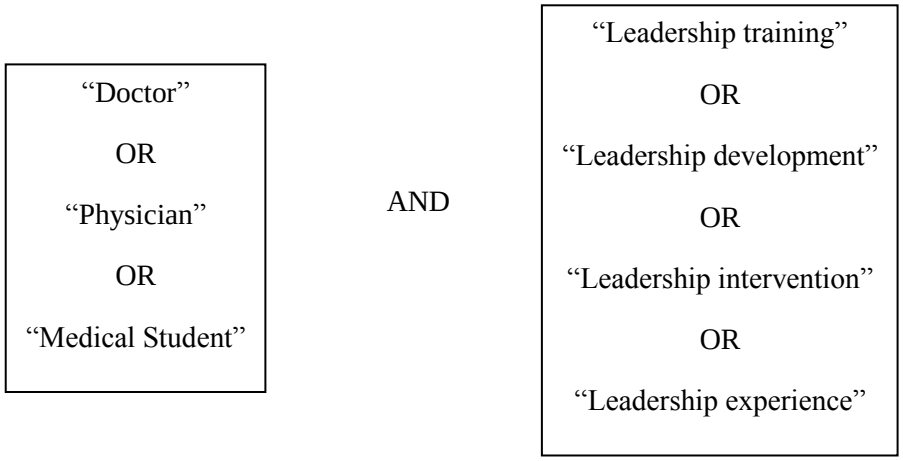


Figure 2.1 Final search terms for systematic review

2.3.2 Inclusion / exclusion criteria

Having finalised the search terms, further work was undertaken to decide upon the scope and strategy for sourcing articles. This work needed to ensure that the searches were as broad as possible, but were also able to focus right down onto studies where there might be useful learning and evidence for the overall project. Key to this was developing and implementing effective inclusion / exclusion criteria. The main inclusion criteria placed emphasis on the presence of:

A leadership-focused intervention - focusing back on the research question "do leadership development interventions improve leadership skills in doctors?" it is clear that all selected publications must have a described intervention related to leadership development. This can at times be very ambiguous as definitions of leadership vary widely and many different broader educational interventions may claim to have an impact on 'leadership'. For this review, the intention was to include studies where leadership was the primary focus of the intervention alongside studies where leadership was a significant component of a wider educational programme. It was decided not to restrict the setting of these leadership interventions, so regardless of whether they were classroom-based or work-based they were included. Studies that solely had a single focus on concepts such a change management, mentoring or quality improvement, rather than a broader leadership approach, were excluded. Academic programmes that exclusively focused on the development of educational leaders were also excluded.

A clear target population – this project was focused on leadership development within doctors and medical students so programmes involving other healthcare professionals were excluded from the mainstream review, although they were read through in case there was applicable learning.

Clear outcomes – each study needed to have at least one well-defined outcome measurement, even if in the end it may not have demonstrated any significant change. Across the different studies this measure could take a number of different forms including use of self-assessment tools, confidence scales or some kind of proxy marker of impact. Studies with no intervention or outcome measure(s), or those where the only focus of an evaluation was on participants' experience, were excluded from the search.

English language - unfortunately non-English citations were excluded due to the cost and logistical implications of having them translated. It has to be acknowledged that although the majority of the leadership development work across healthcare has been published in English, this approach may have led to missing out on some important evidence.

Clear study design - one area where search 'filters' and restrictions were not applied tightly was around the type of study described. Although many systematic medical searches have extensive pre-defined search filters, designed to find the type of study design deemed most suitable for the clinical research question, the aim of this systematic review was to be as inclusive as possible. Medical education research tends to be broadly dispersed across many different methodologies so any mechanism that focused too closely on one particular study design would be likely to detrimentally affect the sensitivity of the search (78). It was important, though, that the description of study design was clear so results could be taken in appropriate context.

2.3.3 Database searches

A total of eight different databases from across the medical, education and social sciences paradigms were interrogated in August 2009 using the search strategy illustrated in Figure 2.3. The five databases:

EMBASE

HEALTH BUSINESS ELITE

HMIC

PubMed (including MEDLINE)

PsycINFO

were explored on a combined search using the facilities available on the National Library for Health (www.library.nhs.uk), which has since changed its name to NHS Evidence. This advanced facility enabled effective management of duplicates and shortened the time that would have been taken in running individual searches within each database. The output from this combined search was independently reviewed by the lead researcher (RK) with assistance from his two junior researchers (RH and AR). This process involved the lead researcher (RK) and subsequently the two junior researchers independently reviewing the titles and abstracts of each of the search 'hits'. A 'traffic light' scoring system was used to assess the studies against the inclusion / exclusion criteria outlined above. Studies with clearly described interventions and outcomes were labelled as 'green' and the studies that clearly had neither were labelled as 'red' and were excluded. Those where it was not clear on a first review whether or not they could be included were given an 'amber' label and were put forward for further review and discussion once the three researchers came together.

At the first discussion stage, the combined 'green' and 'amber' studies were reviewed using the concepts derived from the Cummings review (74), and a team decision was made to decide whether they should be labelled 'red' and excluded, or be passed on to a further level of review, using the full text version of the study (where available). In addition, a 'blue' category was created for studies that clearly did not meet the inclusion criteria, but had a relevant subject matter and might contribute to developing an understanding of the context of this work.

For the second discussion stage full text versions of all of the remaining 'green' and 'amber' papers were sourced; these were found in all except for 4 papers. These papers were re-reviewed by the three researchers and after discussion any papers not using doctors or medical students as subjects, or without clear interventions or outcomes were excluded. Data was extracted using the framework published in the Cummings systematic review (74). The different steps in this whole process are illustrated in Figure 2.4 in the results section.

Three further databases were searched individually by the lead researcher (RK), and any additional studies found were scored using the same traffic light system described above. Due to their other time commitments it was not possible for this search, nor the work described below focused on the grey literature, to also be independently undertaken by the other two researchers. As these databases were less likely to yield important studies than the 5 in the original study this was not anticipated to be a significant problem. These three databases were:

Educational Research Information Centre (ERIC)

Topics in Medical Education (TIMELITT) (79)

Research and Development Resource Base (RDRB)

2.3.4 Grey literature searches

The search of grey literature aimed to be as comprehensive and deep as possible (75), but could not claim to be systematic, and is often described as a 'hand search'. A search using a single web-based search engine (Google) of the terms "leadership training" AND "doctor", which is only part of the search outlined above, generated over 3.3million results. Therefore the work to explore the grey literature focused on six different angles:

a) Exploration of the references at the bottom of each of the selected studies to look for papers that had not been picked up in the original database searches
b) 'Google Scholar' was used to repeat the database searches to look for associate reports, books or papers not listed on the databases above
c) Searches within 'Open Grey', a European collaboration comprising of 700,000 open-access documents across Science, Technology, Biomedical Science, Economics, Social Science and Humanities.
d) Web-based searches of institutions, academic organisations and commercial organisations with well-established credentials in the leadership arena, looking for leadership articles and reports linked to healthcare. *Related links* sections of these websites were explored where available
e) Searches of specific databases / libraries. With the support of their librarian, a search was made of the King's Fund databases to look for any supporting evidence
f) Over the 3 years in which this work was conducted and written up, articles and reports were assimilated through personal contact and discussion with a wide range of experts and stakeholders in leadership development within healthcare.

2.3.5 Search dates

One final factor in the methodology of this systematic review regarded the timing of it. The database searches were undertaken in August 2009, with the grey literature searches conducted on an on-going basis over the following 3 years. This meant that any papers published after early 2009 were not included in the main search.

Although it was not possible to entirely re-do the full systematic review again during the conclusion of the writing up phase, the database searches were each revisited for an overview

scan of the literature that has been published between 2009 and 2012. Any papers deemed to be of significant interest were obtained in full text format and were taken through the 'traffic light' methodology to assess whether or not they should be included. For logistical reasons this was only done by the lead researcher (RK). This review has not been re-run to include literature published more recently than 2012.

2.4 Results

This section describes the results of the systematic review that was undertaken in August 2009. The first sub-section outlines the numbers of papers found at each stage of the systematic review, and the second sub-section presents a more detailed qualitative summary of the learning from the key studies that were found. The subsequent sub-sections present the findings from the other important sources and searches detailed in the methods section above.

2.4.1 Search results by stage

Having followed the systematic review search methodology described in the detailed sections above, 1378 potential studies were found during the initial search of the five leading databases EMBASE, HEALTH BUSINESS ELITE, HMIC, PubMed (including MEDLINE) and PsycINFO.

Figure 2.2 details the numbers of studies found at each stage and illustrates how each successive stage of the review reduced the number of applicable studies down to only those with close relevance to the project.

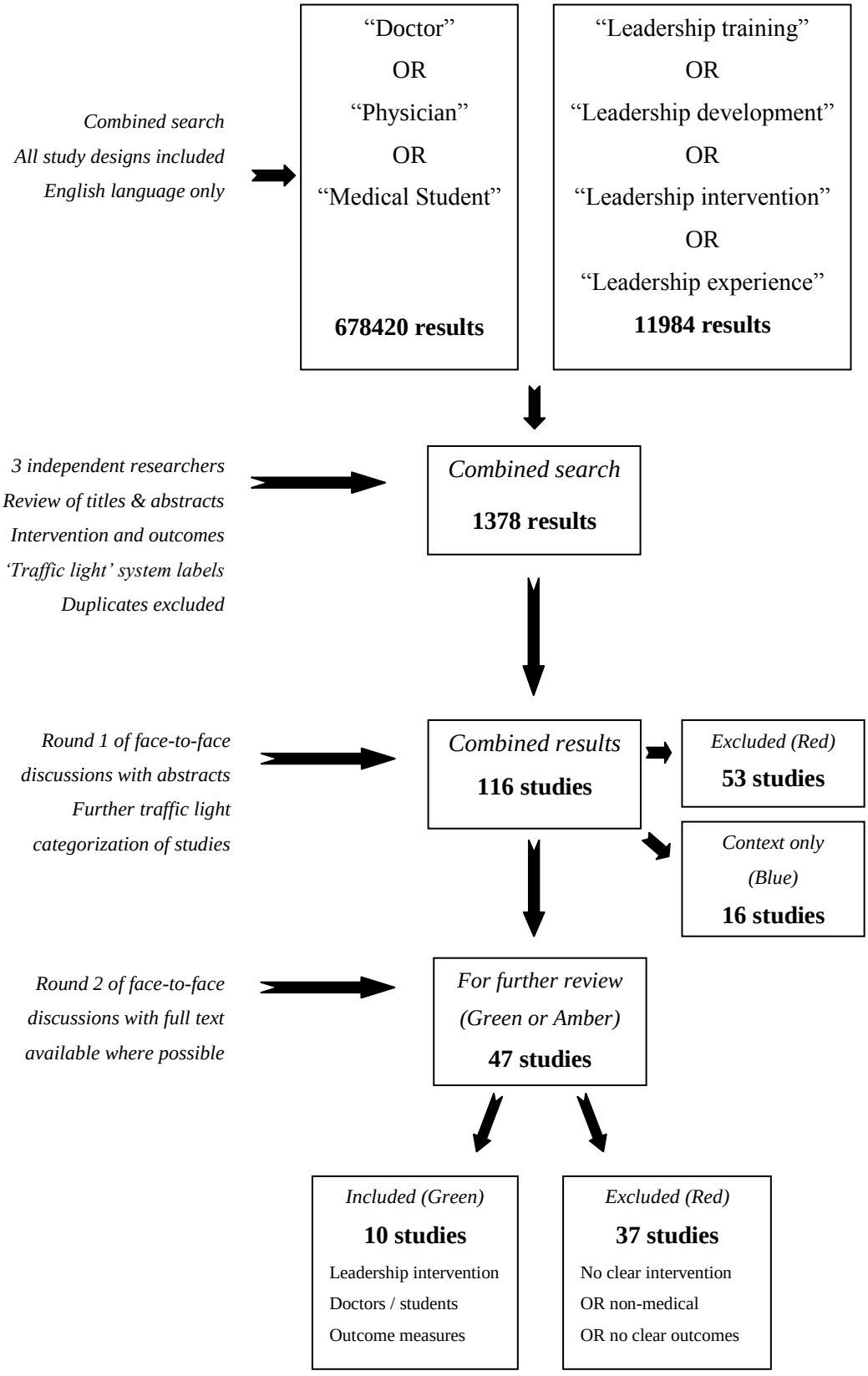


Figure 2.2 **Summary of numbers of papers found at each stage of the main search of the databases EMBASE, HEALTH BUSINESS ELITE, HMIC, PubMed (including MEDLINE) and PsycINFO.**

The detailed descriptions and learning from these 10 included studies are described in sub-section 2.4.2 below. Through careful review of the studies referenced and cited by these 10, a systematic review of teamwork training interventions in medical student and resident education (80) was identified. Scrutiny of the references of this review led on to a further 4 papers of potential interest; these are discussed in more detail in sub-section 2.4.3.

The searches conducted within the additional three databases, Educational Research Information Centre (ERIC), Topics in Medical Education (TIMELITT) and Research and Development Resource Base (RDRB) did not yield any new results. There were a significant number of papers relating to leadership, and in particular focused on leadership development of academic staff but, disappointingly, none met the inclusion criteria for the systematic review.

The searches of the grey literature were more fruitful and, although they did not highlight any further studies of leadership in doctors or medical students with described interventions and outcomes, there were articles and reports retrieved that helped to provide important context to the programme. Some of the most useful documents found are described in sub-section 2.4.6 below.

As described in the methods sub-section 2.3.5 the search of the main 5 databases was re-run in August 2012 in order to gain an overview as to whether other significant studies had been published in the intervening 3 years. Clearly at this stage these studies could not influence the design of the leadership intervention in this research, but they may help to contextualise the results. The depth of the review set out in Figure 2.2 was not repeated in this 2012 re-run so it is acknowledged that some studies may have been missed. Of relevance, in early 2012 *Medical Teacher* published Steinert et al's BEME systematic review: Faculty development initiatives designed to promote leadership in medical education (81). This comprehensive systematic review took a very broad definition of leadership and explored the literature for studies with interventions that promoted leadership within medical education. Whilst some of the results illustrated studies in which leadership was the primary focus of the intervention, many others involved leadership as a component of a wider educational or academic development programme. Detailed examination of this review allowed for cross-referencing of four of the studies that were found also found in the main database search of this study, and the addition of two relevant studies that were not found in the database search (82) (83).

One further paper that fulfilled the original inclusion criteria was found in the re-run of the database search. Juriza et al's 2011 study of outdoor camps for medical students from the Universiti Kebangsaan in Malaysia (84) is described in more detail in sub-section 2.4.5.

2.4.2 Summary of studies that met inclusion criteria

Although there were many detailed descriptions of innovative medical leadership development programmes found within the main database search, the vast majority did not have any recorded evaluation strategy. Where there was an evaluation it tended to be focused on participants' perceptions around the quality of the teaching or facilitation. Most of the studies were not aligned to any curricula or specific learning need and only a very small number had made an attempt to measure any outcomes. Where there was an evaluation the focus was on assessing changes in the confidence of the participants over time.

Ten studies found from the original database search, and a further two as references from Steinert et al's systematic review (81), were deemed to have met the inclusion criteria, although as discussed below, the extent to which their focus was around leadership competencies, was not always clear.

The majority of the included studies (7 out of 12) related to short courses, programmes or workshops of between 1 and 5 days duration. Evans et al described an outdoor-based leadership training intervention for family practice interns, which aimed to improve group problem solving, team building and communication skills (85). This intervention was incorporated into 2 out of 13 orientation (induction) programs within the University of Washington Family Practice Residency Network, with the other 11 programs acting as controls. A survey was developed to assess perceptions of trust, group awareness, effectiveness and problem solving within the intern class one month after orientation. This demonstrated some statistically significant improvements in the post-intervention survey results. Stoller et al also focused on teambuilding and leadership in a 1-day retreat for first year residents in internal medicine (86). Using a combination of taught content, discussions and experiential exercises participants demonstrated increased self-assessment scores in a pre- and post-intervention evaluation survey. Both of these interventions were brief in duration, and the post-intervention evaluation taken shortly after, so it can be argued that it is difficult to meaningfully assess the longer term effect of these interventions.

Two of the studies describe ‘short courses’ which were conducted in the undergraduate arena. One involved a week-long retreat for 23 students from across the USA, with the aim of developing leadership through the execution of a medical education project as well as creating a national community of student leaders in medical education (87). The second focused on a student-selected component (SSC) at Hull York Medical School, UK involving 20 students (88). Both studies showed self-perceived improvements in confidence and ability in the students and coherently argue the case for leadership development starting at an early age.

All four of these studies illustrate ‘short courses’ that were evaluated using pre- and post-intervention self-assessment surveys. The questions used were predominantly focused on Kirkpatrick Level 1 (reaction / satisfaction) and Level 2 (self-reported learning / knowledge) evaluation methodologies (89) and there was no longer term follow-up conducted. The short term focus of both the intervention and the evaluation are significant barriers to using the learning to shape this research.

Three further programmes, that were based around the ‘short course’ model of delivery, made an attempt to follow-up attendees to see whether there were any measurable long term benefits of participating in a particular leadership programme. In the first of these studies, Leslie et al designed a 3-day programme to develop young paediatric leaders in the USA (90). Participants, who were under 40 years of age, and who had less than 5 years in clinical practice, were required to complete a detailed needs assessment, attend the 3-day training programme, and commit to one leadership related behaviour change being implemented over the subsequent six months. The quantitative evaluation strategy used pre-test, post-test and delayed post-test questionnaires to self-assess across 20 leadership competencies and showed improvement in many domains immediately following the programme, and that for the most part it was sustained 6 months later. Although this study has introduced a delayed post-test evaluation, which might describe any longer term learning, there was no measure made of any more objective changes, nor their potential influence on the organisations in which the doctors were working.

In the second study of a ‘short course’ programme with longer term follow up, Steinert et al describe a 2-day workshop aimed at developing the leadership, management and administrative skills of 20 members of the clinical faculty (91). The content of this initiative was focused on self-management strategies, leadership styles and skills. In addition to an immediate post-workshop evaluation questionnaire half of the candidates participated in a semi-structured telephone interviews around 1 year later. These interviews revealed that over the 12 months they had developed strategies around setting short-term goals, protecting time,

setting agendas and handling paper more effectively. With other behaviours such as delegating, saying "no" and adopting different leadership styles there had been much less success. The authors concluded that if sustained improvement was to be achieved that these areas would need follow-up and reinforcement. This introduces some interesting ideas that might support the development of a work-based intervention where there is ongoing reinforcement of learning at a number of stages through the programme.

In the third study of this type Duda et al describe a one-day course that was run to introduce junior academics to concepts around leadership (82). This paper was found through scrutiny of the references in the Steinert 2012 systematic review (81), and is of interest as it describes an evaluation methodology that sought to use 'motivation to seek a leadership position', and 'actual academic career progression 3-4 years after the course' as outcome measures. Respondents reported an increase in motivation to consider leadership positions, and to apply the skills learned in the course. Despite the conclusions of the author, to what extent this delayed finding can be attributed to a one-day course seems to be open for debate. Although this study was only found in the 2012 re-run, the outcomes used in this study might be helpful in shaping future long-term follow-up.

Five of the studies took a different approach to the leadership development intervention, and describe programmes that were between 6 months and 2 years in length. Hemmer et al report on 1-year long postgraduate study which was focused on developing leadership and management skills in residents and fellows in laboratory medicine and pathology at Mayo Clinic, Rochester (92). This formal programme used lectures, interactive sessions, case-scenarios, team building exercises and team presentations to cover a well-defined curriculum. Trainees were excused clinical duties at this time. Participants were assessed using a pre- and post-course knowledge based assessment, and through their performance in a team-based project presentation. These demonstrated an improved knowledge of leadership and management issues immediately following the programme. It is of note that this was the only study of those that met the inclusion criteria to use an outcome measurement tool that was not based on the principles of self-assessment. This opens up some interesting critique that is picked up in detail within the sub-section on self-assessment in Chapter 4.

Another programme, open to paediatric interns, that was closely aligned to a leadership curriculum, was the 2-year long physician leadership programme for 28 physicians, established by McAlearney et al at Columbus Children's Hospital, Ohio (93). The programme consisted of regular interactive presentations aimed at building organisational awareness and basic core

knowledge around leadership. Evaluation once again was through a pre- and post-intervention self-assessment focused on Kirkpatrick levels 1 and 2 (89). In addition participants were surveyed one year after completion of the programme to ask about their perceptions of the programme and of their opportunities to apply the skills they learned. Comments suggested that participants' perceived that the programme had changed their leadership behaviours, particularly in the areas of decision making, conflict resolution, business planning and managing people. This study, with its 2 year programme and paediatric setting, was influential in supporting the development of the intervention in this study. From what can be gleaned from the paper the focus was on improving physicians' leadership skills and increasing the understanding of the strategic goals and direction of the organization in which it was set.

Osborn and DeWitt describe the development of a faculty development programme focused on the development of leadership in medical education (83). As with the Leslie (90) and McAlearney (93) studies this had paediatrician participants, although the emphasis here was on developing academic paediatric leaders who could institute the changes needed for medical schools. The intervention was workshop-focused with a number of events run over a 12 month period. The evaluation methodology involved a retrospective pre-test and immediate post-test survey that suggested gains in knowledge. The 2-year delayed post-test questionnaire pointed towards an increase in leadership positions, promotions and grants, although it was acknowledged that this was unlikely to be a directly attributable effect, and there was no control group. There were also outcome measures in terms of the numbers of workshops the original participants went on to deliver, and the number of attendees, which gives a crude assessment of one aspect of ongoing impact of this intervention.

Morahan et al's description of a 6-month part time fellowship also focused on developing the leadership skills of academic medical faculty within the context of Academic Health Centres (94). The programme was held off-site and consisted of around 10 days of learning from lectures, assignments and small-group work over a 6-7 month period. The programme was externally evaluated using a mixed methods approach. This consisted of pre- and post-intervention surveys that explored participants' perceptions of their knowledge and skills in dealing with the wider system, and these findings were triangulated by qualitative in-depth interviews with participants and organisers at the end of the programme. Some of the key recommendations that came from this paper were to focus on specific skills that could be learned, link learning to real work situations, involve senior leaders in the running of the programme, encourage reflection and development of self-awareness and use the leadership training to develop networks. This study provided helpful input, both into the thinking around the design of the intervention, but also in developing an evaluation approach. One of the key

stated lessons from the work was that time and expertise for evaluation needs to be provided from the very beginning of planning a programme, which resonates strongly with the planning and development of this project.

The final study to meet the inclusion criteria was Dannels et al's work on the 'Executive Leadership in Academic Medicine (ELAM)' programme (95). This study, which explored a year-long leadership development programme for senior women faculty members in US and Canadian medical schools, took a different approach to the evaluation methodology. As with many of the other studies described above, a pre- and post-intervention methodology was adopted. However, in addition to the study cohort, control groups were established. On one occasion they came from women who applied, but were not accepted, into the ELAM programme, and on a second occasion were women from the Association of American Medical Colleges. The results showed significant improvements, compared to the controls, with 12 different indicators. The difficulty with this is that the comparison groups cannot be considered to be a 'case-controlled situation' in the same way as is used in clinical studies, but the authors' conclusions that the "inherent limitations of leadership evaluation studies emphasises the importance of using multiple sources of data to establish outcomes" provide useful advice in shaping the design of this research (95).

2.4.3 Summary of studies found via citation and reference

One of the papers that was found during the searches was a systematic review of teamwork training interventions in medical student and resident education (Chakraborti et al. 2008). Although this review, which describes 13 studies of curricula designed to teach teamwork to medical students, had a main focus on team-working, rather than leadership development, on closer examination six of the programmes also incorporated some aspects of leadership training.

Two of the six programmes, described by Stoller and Evans respectively, were also found individually through the systematic search (85)(86). Of note none of the studies used a particularly work-based strategy to develop leadership competencies, and instead used brief classroom-based interventions. The two postgraduate programmes took doctors out of their working environment for just a single day to develop leadership skills (96)(97). The concern around this approach, and the reason that it has not been adopted for this study, is that the intervention is unlikely to have enough depth to lead to a long-term impact. One of the undergraduate studies took an approach that tried to align learning to a clerkship over an eight-

week period (98) whilst the other undergraduate programme used workshops and e-learning materials over a one-year period (99).

In terms of evaluation methodology, although many of the studies described in the review used validated instruments to assess teamwork, evaluation methodologies were generally weak, and tended to be focused on reaction to the experience and self-assessments of knowledge. As with the vast majority of work in this field no studies used randomisation and only a few had any control group within the study design. There was also a lack of long-term follow-up of learner outcomes that further limited the conclusions that could be drawn. The question around a control group and thinking hard about the logistical possibilities around long-term follow up were important considerations in the design of the methodology within the research. This is discussed in more detail in Chapters 3,4 and 5.

2.4.4 Summary of studies which added context

As highlighted in Figure 2.4 there were a number of studies that did not meet the inclusion criteria, as they did not have both a well-defined intervention and an evaluation strategy. However, in reading through their full text, ideas and concepts that could help to develop a deeper understanding of the context of this work were extracted.

In the first of these papers McAlearney (100) linked up a number of leadership development programmes with quality improvement initiatives. A qualitative evaluation highlighted the potential synergy between leadership learning and quality improvement, and indicated how organisations could use a leadership development approach to drive definitive improvements in quality and efficiency. With the NHS seeking to make between £15-20 billion of efficiency savings between 2011 and 2014 (101), it is undoubtedly a strategic advantage for leadership development programmes to have a secondary outcome of improving efficiency, so this is something that it feels helpful to consider incorporating into the intervention design.

Beecham et al (102) describe a 6-month long leadership development intervention for general practitioners (GPs) that was built around an e-mail and telephone coaching approach. Although not formally evaluated the authors described how this intervention gave GPs a ‘safe place’ within which to explore their individual personal development. Another approach to leadership development comes through ‘understanding oneself’, and a number of programmes describe the use of the Myers-Briggs Type Indicator (MBTI) (103) as a way of introducing leadership competencies to doctors in training (104). This tool was reported to have allowed participants

to openly explore reactions, describe personal examples and gain a deep insight into interpersonal interactions that have worked or not worked. This is a well validated and widely used tool, and from experience is a very powerful way to explore some of the concepts around self-awareness, managing yourself, building and maintaining relationships and managing people, all of which can be found in the first 3 domains of the MLCF (19).

Wahlstrom and Sanden, working in the context of an orthopaedic department in Linkoping, Sweden recognised that an increasing proportion of healthcare delivery involves teamwork and cooperation between various occupational categories (105). They utilised a training ward to develop a 2-week long simulated scenario-based programme that was attended by over 450 students from a number of different professional groups. Other studies focused on developing the leadership capabilities of faculty in order to then more broadly build leadership capacity across the system. Hill and Stephens describe a programme established at the School of Medicine, University of Southampton that takes a whole-systems approach to leadership development (106). The aim of the programme was to enhance and empower the course co-ordinators across the medical school, and through this to create an environment where leadership learning was rich. Although neither of these reports had strong evaluation approaches these ideas around team-working and systems leadership are important factors that influenced the thinking around the leadership intervention in this study.

Two final papers placed their focus on how students and doctors could gain leadership competencies through participation in postgraduate degree programmes. Crites et al (107) describe the initial development of a five-year programme where an MD degree, leadership curriculum and an MBA or MPH degree are delivered concurrently. The programme was developed using a systematic qualitative inquiry approach, and this resulted in an innovative curriculum and programme of learning. Unfortunately there is no published evaluation or review of the impact of the programme, so it is not possible to critique the impact. Schwartz et al (108) describe a number of similar programmes designed to support physicians to matriculate through a series of educational steps, which may eventually culminate in a masters-level degree in health administration (MHA) or business administration (MBA). They conclude that physician leadership training should be local, offered as a long-term package and be physician led, stating "the concept of physician leadership will not and should not be taken seriously by non-physician health care executives until the physician community becomes as serious about leadership and management training as it is about clinical training." Although of interest, neither of these papers had any direct influence on the design of the intervention in this research.

2.4.5 Summary of studies found in August 2012 update

As explained in sub-section 2.3.5, the original database searches within the systematic review were each revisited in 2012 for an overview scan of the literature that has been published since 2009.The main paper that emerged, Steinert et al's 2012 BEME systematic review, 'Faculty development initiatives designed to promote leadership in medical education'(81) is described in sub-section 2.4.1 above, and in the discussion section 2.6 below. There was only one new study that would have met the original search criteria that was picked up in the repeat searches. In this paper, Juriza et al describe a series of three residential 'camps' for medical students, placed throughout their 5 year course, and designed to bring on their team working and leadership capabilities (84). As with many of the previously described studies, the evaluation was undertaken by self-assessment with a mid- and post-intervention survey, which in this case focused on participants' perceptions as to whether they had achieved the programme's learning outcomes. The paper's discussion focused on the reported value of 'experiential learning' (109) and how the camps had provided a "relaxed and friendly ambience" within which to learn and develop. Although there are some inherent weaknesses with this paper, with no control group and a somewhat vague description of the leadership intervention, the discussions around experiential learning have relevance to the discussions in the final chapter of this thesis.

In one other study Gowan (110) reports on a pilot leadership development for postgraduate trainees in a large UK teaching hospital. The programme consisted of 5 one-day modules held over an eight month period and, although not formally evaluated, feedback from participants was generated. Perhaps the most illuminating part of this, on the background of positivity about their experiences, was that the 360-degree feedback tool was "particularly disliked by participants, dismissed by many of the assessors and did not add any value to the process". The small size of this study means that even if it had been found in the original 2009 literature search, it may not have influenced the thinking around 360-degree feedback tools, but it does provide some strong scepticism that can be fed into discussions in Chapter 8.

One other important paper was found in the repeat search. Coltart and colleagues' Lancet paper describes the background to and basis of the UK's Chief Medical Officer's Clinical Advisors Scheme (111), which takes a very strong work-place based experiential approach to learning. In this scheme, clinical advisors, at the level of junior doctors, are seconded from their clinical placements to work for one-year full-time within one of a number of high-level health organisations. The advisors worked alongside senior colleagues to develop skills in project management, research, policy development, team working, analytical evaluation and leadership. The scheme has so far been able to place 50 junior doctors in these roles and from the feedback and career pathways described in the paper appears to have had a significant

impact on many of the individuals involved. Unfortunately no evaluation was undertaken, so it can really only be taken as being the reflections of a cohort of the individuals involved in the programme.

2.4.6 Summary of studies from the grey literature

In sub-section 2.3.4 the approach taken to searching the grey literature was described. A large number of papers, articles, reports and websites were reviewed, and through this reading, and a number of conversations with different experts in this field, further ideas about leadership development interventions, and how they might be evaluated were gained. A small number of the resources that were reviewed are described below in order to demonstrate some examples of the different types of resources that were found.

Two well-written reports, by Mountford and Webb, and published within the healthcare consultancy firm McKinsey & Company, provide strong case-study based evidence from across the world to suggest that for health care systems to unlock high performance and to be able to transform themselves, they need to harness the energy of their clinicians to be organisational leaders (37) (36). The authors describe what clinical leadership is, explain why it works and explore underlying barriers to clinical leadership within the healthcare system. A significant proportion of the writing is devoted to describing practical steps to build stronger clinical leadership. These steps, derived from the case-studies described, are based firstly around the need to create an enabling organisational and external environment for clinical leadership, where organisational performance matters and where any disincentives for clinicians to lead are removed. The second component focuses on the need to develop outstanding leadership development programmes – their suggestion that these should be 'longitudinal', work-based and experiential resonated strongly with the emerging ideas for the shape of the leadership development intervention in this project. Finally, the authors suggest that perhaps the toughest barrier to clinical leadership can be seen in "the historical habits and beliefs of clinicians themselves". They conclude that any successful programme must explicitly address these mindsets. They suggest that this is best achieved through the bringing together of clinical leaders "in a community of the like-minded where they can share information and stories and realise they are not alone in the world." It is the essence of this 'community' that is a concept that it feels important to try to create through this work.

Another excellent source of background reading came from the extensive work undertaken in medical leadership by the NHS Institute for Innovation and Improvement, in conjunction with the Academy of Medical Royal Colleges. This collaboration planned, designed and

implemented the Medical Leadership Competency Framework (MLCF) (19) and through participating as a clinical advisor to the team leading this work, the lead researcher (RK) was able to gain some extremely valuable insights into the vision and strategies around the development and spread of this work. Two important literature reviews, undertaken by Helen Dickinson and Chris Ham from the University of Birmingham's Health Services Management Centre, provided the international evidence that underpinned much of this work (112) (45). This evidence suggested that through engaging doctors with change processes, improvements in performance may be achieved. Looking at the source materials described, amongst the countries reviewed, Denmark seems to stand out for its efforts to engage clinicians in leadership roles and to give them training and support. There are also interesting case studies described, such as that of Kaiser Permanente (41), in the United States, where the close alignment of the health plan and the medical teams contributes significantly to the high levels of performance that characterise the organisation. The NHS has a good opportunity to learn from these international experiences and become an exemplar in healthcare that is characterised by strong clinical leadership.

Although the focus of the systematic review was around medical leadership, the searches of the grey literature led to interesting resources and material from outside the healthcare sector. One of the most helpful reports focused on the strategic principles around leadership learning that were applied to the establishment of the National College for School Leadership in the UK, which was founded to support the leadership development of school leaders and head teachers (113). In this work, Stoll describes the importance of leadership learning being a coherent experience, connected to work-based experiences where both learning and impact will be at their highest. There is strong emphasis on leaders taking responsibility for their own learning and displaying a commitment to life-long learning (114). The report also highlighted the 'distributed leadership' or 'shared leadership' approach taken within schools, that with the introduction of the MLCF (19), has become one of the key principles underpinning medical leadership development in the UK. Although not medical in setting, these principles are helpful in providing guidance to the sort of work-based approach that could be taken in the study.

Perhaps the largest source of leadership literature comes from business and industry, so it was difficult to gain anything more than an overview of what specific learning might apply to leadership development within the NHS. Fitting in with the trait theories discussed in sub-section 1.4.2 a significant proportion of this literature focuses on the leadership 'traits' seen in the companies' top leaders. However, there is increasingly a move towards thinking more about establishing leadership development programmes for wider groups of employees. Many

of the most successful FTSE 100 index companies in the UK (the 100 companies in the UK with the highest market capitalisation, which together represent around 81% of the value of the London Stock Exchange) have placed significant investment in their in-house leadership development programmes. An example of this would be Diageo. Following up on background reading, through personal contact with the Head of Leadership, the lead researcher (RK) was able to meet with a number of individuals within Diageo (in 2012 the 12th largest company on the London Stock Exchange) to explore their approach to leadership development within the company. They had created a year-long experiential leadership programme, then in its second year, that had been made available to 900 of the company's top leaders across the world. The key investment was in the external performance coaching that underpinned the programme. These 'coaching conversations' allowed participants space and opportunities to explore their personal values, to be challenged and to grow. The outcomes chosen to measure the success of the programme were heavyweight; staff satisfaction, staff retention and ultimately overall financial performance of the company. Although these are the sorts to outcomes that will be extremely hard to attribute to a single programme, this does provide an important reminder that the outcomes of leadership development programmes can be wide-reaching.

2.5 Conclusions

The opening objective of this study was to perform a focused systematic review of the medical leadership literature, and to analyse key learning and themes around medical leadership development initiatives. Despite a comprehensive search, involving a number of educational sources and databases in addition to the more commonly used medical databases, only 12 studies were found that met the inclusion criteria. Most of these studies were short-programmes of face-to-face training, but there were some that had a more experiential work-based focus and that were run over many months. Almost all of the studies were underpinned by an evaluation strategy that compared pre- and post-survey self-assessment data, although a minority of studies adopted other strategies, such as the use of qualitative research methodologies. Very few utilised control groups.

The themes and learning from the studies found have been helpful in the design and planning of the methodology behind this project. In addition, there were also several opportunities to gain learning from some of the studies that did not meet the inclusion criteria, and from the grey literature. It must be acknowledged, however, that it is likely that the more recently published Steinart review (81) would have further influenced the design of the intervention had it been available during the planning phase.

2.6 Discussion

This review of the published literature has provided many different useful inputs into the design and direction of the leadership development programme in this project, which relates back to the first and second objectives of the thesis. It has highlighted that there have been relatively few leadership development programmes for doctors previously described in the literature. A number of the studies reiterated the importance of aligning learning objectives to a leadership curriculum, in our case the MLCF (19), and helped in the difficult process of developing effective assessment tools or evaluation methodologies that would be able to measure leadership outcomes.

In order to achieve both breadth and depth with this systematic review it was important that the methodological design of the searches was as strong as possible. As described in sub-section 2.1.2 a specific approach had to be taken in order to be able maximise the chances of finding applicable studies within the education, leadership or social sciences literature, as well as more mainstream 'medical' databases. It is important to acknowledge that the use of 'doctor / physician / medical student' as a search term in effect restricted the systematic review to the studies focused on *medical* leadership, and this may have resulted in important evidence from other non-medical clinical groups being missed.

The principles that the two BEME papers (73)(78) offered helped to enablee the development of a methodological framework that could be closely followed. The efforts to review lists of references, to explore studies that cited the retrieved papers, to revisit the searches with an August 2012 update and to explore the grey literature all added to the strength of this work. However, as with any search of this nature the difficulty in achieving full coverage has to be acknowledged. Not all databases or possible search terms were covered and so despite the rigorous approach taken in this review, there may have been key studies that were not found.

Of the papers that were found to meet the inclusion criteria, the majority described studies where the leadership intervention was a form of short-course, usually held outside the context of work. The five studies that had a longer-term experiential approach to the leadership development intervention (83)(92)(93)(94)(95) were of particular relevance to this project, as they helped shape the work-place based nature of the intervention. This was an important aspect of the second objective of the overall project. The conclusions from these studies around the importance of linking learning to real work situations, involving senior leaders in the running of the programme and encouraging reflection and development of self-awareness were extremely useful points that were built into the design of the intervention. There was also

an emphasis on the value of using a leadership training programme to develop professional networks. The difficulty with these studies is that the level of detail described in each of them made it hard to gain anything more than some outline principles in terms of shaping the nature of the intervention used in this research.

Perhaps the most helpful lesson from the published studies was that a comprehensive evaluation strategy needs to be developed from the very beginning of planning a leadership development programme. All of the included studies relied on self-assessment surveys of participants. These were predominantly focused on exploring participants' reactions, but in the case of the Hemmer study (92) for example, the survey was a knowledge test. Most of the studies ran these surveys pre- and immediately post-intervention, but in a minority of the studies participants were asked to complete a post-intervention survey months, or even years, after the leadership development intervention had been completed. Some of the studies took a variety of approaches to augment this self-assessment data with some qualitative evaluation data, obtained through a number of different methods. Within the discussion sections of some of these papers the potential to evaluate the impact of the leadership development intervention using a more objective assessment, based on actual performance rather than just self-assessment, was explored, but none of the papers actually did this. The idea of randomising participants to a control or study group was also considered in one of the papers. However, despite the recognition of the potential advantages of these potentially more robust methodologies, the papers' authors concluded that the logistical difficulties and inherent biases of these approaches would outweigh any methodological advantage of the increased objectivity. The learning from these studies and how it has been used to design the evaluation methodology for this study is described in sub-section 3.3.7.

Steinert et al's 2012 BEME systematic review (81) came too late to be able to influence the design of the intervention – it was only picked up in the August 2012 search update – but has been useful in some of the interpretation and contextualisation of the results. The key learning from this review was around highlighting the potential impact of a work-based leadership development intervention, and also in emphasising the potential value of a longer term evaluation (3-4 years post intervention) that made some assessment of medium to long term impact of the leadership development intervention. This longer term approach may have had significant value in this study, a discussion that is picked up in the final chapter.

Finally, there were a number of other specific points that came from the different studies mentioned above. Participants in many of the programmes recognised the value of being given

some space, and the opportunity for self-directed learning, within their programme (90)(91)(93). The development of self-awareness was a key component for many of the leadership interventions, and in more than one study a framework for this had been created through the use of MBTI assessments (104). This provides helpful evidence to support the 'developing personal qualities' domain of the MLCF. Interestingly, the idea of using a specifically designed 360 degree multi-source feedback tool was not supported by the one group of participants where it was trialled (110). Although these views are important, the numbers in this single study are small, so this finding does not in itself need to rule out the use of 360 feedback tools.

Further detail and discussion around the development of the intervention can be found in Chapter 3.

Chapter 3 – Developing the intervention and the research design

Chapter Overview

This chapter focuses on the development of the leadership intervention and outlines the design of the research methodology used to evaluate it. The chapter describes the vision of a trainee committee within the London Speciality School of Paediatrics and Child Health, and discusses how their work, and the experiences gained, underpinned a work-based leadership development programme. The chapter details the selection of the trainee committee cohort, and discusses the design, refinement and implementation of the leadership intervention itself. It also describes the trial of a leadership multi-source feedback tool which aimed to support the learning and development of the participants who wanted to use it. The chapter also provides an overview of the mixed methods research methodology used in this thesis, and describes the technical aspects of this and the rationale behind it.

3.1 Introduction

This chapter builds on the leadership theories described in chapter 1, and the empirical evidence from the systematic review of the medical leadership literature in chapter 2, to explore the selection of the trainee committee cohort, and the refinement of the leadership development intervention. It highlights how this relates to the Medical Leadership Competency Framework. It also provides an introduction to the mixed methods research methodology used in this thesis.

3.1.1 Background to London Speciality School of Paediatrics

The London Specialty School of Paediatrics was established in 2008 as part of a significant national change in the provision of postgraduate medical education in England. Since the advent of 14 Deaneries within England, the responsibility for the organisation, delivery and management of training had sat with these bodies, and there was concern that the relationship between the Deaneries and the relevant Medical Royal Colleges was becoming increasingly distant. In an attempt to address this, the Specialty Schools made joint Royal College and Deanery appointments to their Head of School role and joint processes of governance and accountability were established.

In London, which has around 1000 paediatric trainees on training programmes at any one time, many of the existing Training Programme Directors stayed on within the London Specialty School of Paediatrics, and so in many ways the transition was very stable. However, the birth of the Specialty School, and the appointment of a very innovative and dynamic leader as Head of School, presented opportunities to look at certain key aspects of training in a very different way. One of these areas was around trainee representation and their involvement in the organisation and management of training.

3.1.2 Background to trainee representation in paediatrics

Prior to the establishment of the London Speciality School of Paediatrics, there were two elected trainee representatives on the London Specialty Training Committee (STC) for Paediatrics, and other specialties had very similar arrangements. These roles were taken up by self-nomination and then election. However, with only ever somewhere between 5% and 10% of trainees voting this role never carried any weight. More significant was that there was no structured role of the trainees on the committee, and where the two appointed trainees were on-call or working nights, there would be no representation at all.

The establishment of the London Specialty School of Paediatrics opened up the opportunity to take a new approach. The aim was to significantly expand the size, remit and influence of the trainee committee, to encourage more junior trainees to become involved alongside more experienced colleagues, and to support the development of this cohort with a bespoke experiential leadership development programme.

3.1.3 Leadership learning in paediatrics and the rationale for work-based learning

As with most specialty training in the UK, the historic approach to leadership development within paediatrics has been fairly minimalist. Postgraduate curricula and training programmes have firmly focused on supporting the development of clinical competencies related to delivering high quality care for each individual patient who presents to our paediatric services (115). My personal observations are that focus on developing systems thinking, learning how to improve services and developing as a leader have previously been considered as things to learn only once a paediatrician has reached consultant level. The usual response to this has been for trainees to attend a classroom-based leadership course in the weeks before their consultant interview, so that the ‘leadership’ box is ticked, and for most trainees this has almost always been good enough to get through a consultant interview.

As described in section 1.2 of the opening chapter, and highlighted in the systematic review in chapter 2, there is an increasing recognition within the NHS of the importance of clinical leadership in developing and improving services. Although there are discussions about how early in medical school these concepts should be introduced, it is increasingly acknowledged that opportunities for leadership learning should be presented to all medical students and doctors throughout their undergraduate and postgraduate training (55)(116). What is less clear is how these learning opportunities should best be created, although a number of the studies described in the systematic review in chapter 2 begin to give some helpful pointers.

The establishment of a trainee committee within the new Specialty School provided the opportunity to design and evaluate a leadership development programme for this cohort of trainees. In planning the establishment of the trainee committee I recognised that a significant proportion of the potential learning and experience gained by the trainees would arise from the work they undertake as part of this role, and that if allowed to involve themselves in all of the work of the Specialty School, and if well supported and facilitated, would open up an outstanding opportunity for high quality experiential learning (109). The theoretical frameworks provided by the transformational models of leadership, where the focus is on valuing actions that enable and empower others (21), and those described by Mintzberg (25) where frontline staff have a strong influence on the direction of travel of the organisation, give important guidance to how this committee, and the leadership development around it, could be set up. As described later in this chapter the concepts of shared leadership (30)(117)(31), described in sub-section 1.1.6, also gave important shape to the leadership development design. This approach also ties in well with the logistical difficulties that many trainees would have in taking time 'out of programme' (118) in order to pause their clinical training so that they could participate a nearer full-time learning programme.

Reflecting back on the systematic review in chapter 2, there were relatively few medical leadership initiatives with work-based components that were found, but those that did further supported the idea of developing this type of leadership development initiative. Morahan's study made a specific recommendation to link learning to real work situations (94), while Beecham et al's study, that had a self-reported positive impact on the leadership development of general practitioners (102), used an e-mail and telephone coaching approach to support work-based learning and development. Hemmer et al's study (92) reported on a 1-year long study that gave residents and fellows in laboratory medicine and pathology the opportunity to learn from a number of different activities, supported by an over-arching curriculum, alongside

their existing work. McAlearney's programme took a similar approach, this time over 2-years and open to paediatric interns (93). Feedback from participants was that it gave them the space and opportunity to learn about leadership and then to make changes to their approach and behaviour in the workplace. Specifically they felt this helped in areas of decision making, conflict resolution, business planning and managing people. These are some of the competencies also found in the MCLF (19) that the leadership development initiative in this study would need to cover.

At the time of developing this work-based leadership development initiative in 2008, this approach could reasonably be described as being innovative. Since then a number of other schemes have adopted this type of approach. The 'Darzi Fellows in Clinical Leadership' programme was introduced in London in 2009 (119), the Chief Medical Officer's Clinical Advisor Scheme widened access to more trainees and became the NHS Medical Director's Clinical Fellow Scheme in 2011 (111) and programmes such as Paired Learning were established (120). These, and other similar regional programmes across different parts of the UK, have given a larger number of medical trainees work-based leadership learning and experience and continue to run and expand. With none of these schemes in place when this project was launched, the aim of designing a comprehensive evaluation strategy was to assess the value of this new approach to leadership development. If shown to be successful, this would help inform further work to extend this type of work-based leadership learning more widely across the NHS. This is discussed further in subsection 8.3.3.

3.2 Objectives

The key objective of this chapter of the thesis focuses on the second of the overall project objectives detailed in sub-section 1.2.3, namely to:

- Design and implement an innovative leadership development initiative to support the learning and development of a newly established postgraduate trainee committee, and design a mixed methods research strategy to evaluate it.

3.3 Methods

This methods section outlines the key principles of the leadership development intervention, and describes the processes involved in the establishment of the trainee committee cohort. It also details the planning and design of each of the four components of the leadership development intervention. It concludes with a final sub-section (3.3.8) that explores the design

of and rationale for the mixed methods research methodology used to evaluate the impact of the leadership intervention.

3.3.1 Outline of the leadership development intervention

The leadership development intervention in this project consisted of four different components, which are illustrated in Figure 3.1. The first three of these could be considered core components, with the fourth one, the leadership 360, as an optional extra. The five domains of the Medical Leadership Competency Framework (MLCF) (19), and the Medical Leadership Curriculum (56) that followed from it were used to ensure a broad coverage of leadership learning was provided.

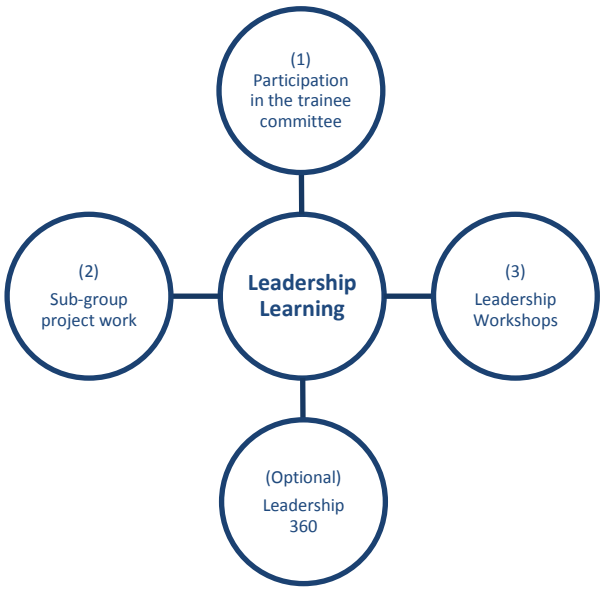


Figure 3.1 Outline of the 4 components of the leadership development intervention

Each of these four components of leadership learning, and the process by which trainees applied to the trainee committee, is described in more detail in the following sub-sections below:

- Sub-section 3.3.3 Application to the trainee committee
- Sub-section 3.3.4 Participation in the trainee committee
- Sub-section 3.3.5 Sub-group project work
- Sub-section 3.3.6 Leadership workshops
- Sub-section 3.3.7 Leadership 360 assessment tool (optional)

The design of this leadership development intervention, like the MLCF, was strongly influenced by Pearce and Conger's work on shared leadership (30), which provided a theoretical framework for the intervention. The paediatricians who were on the trainee committee were all working in environments, whether clinically in their hospitals, or educationally within the School of Paediatrics, where a multiprofessional team approach to leadership, shared across the wider team is prevalent. This strengthened the rationale for developing a work-based leadership development. If trainees had daily exposure to working environments where there was strong role-modelling of many of the leadership competencies described in the MLCF, it could be argued that this would provide a positive learning environment. For that reason participation in the trainee committee, and the opportunities this brought in terms of working alongside Training Programme Directors and the Specialty School Executive Team, became a significant part of the leadership intervention. It was hoped that focusing on project work would also give trainees the opportunity to experience and develop many of the competencies that underpin a shared leadership approach.

In a similar way, Mintzberg's ideas around a healthcare delivery being a professional bureaucracy (25) resonate with the ethos underpinning the trainee committee, and the opportunities given to committee members by senior leaders within the Specialty School. It was hoped that a leadership development intervention based around the work of the trainee committee would give the participants the opportunity to experience and reflect on followership, but also to recognise their important role as leaders within the healthcare system, despite still being labelled as 'junior' doctors. As described below specific workshops were offered in order to cover additional areas of learning need from within the MLCF (19). This described in more detail in sub-section 3.3.6 below.

3.3.2 Key principles of the leadership development intervention

The processes behind the establishment of the trainee committee and the refinement of the leadership development intervention were based around five key principles that were designed from the outset of the establishment of the programme, and that were fully supported by the Head of School (HC) and her executive team. These five principles, that were derived from the theoretical frameworks in the leadership literature described in chapter 1, and in the section above (19)(25)(30), were that:

a) The members of the trainee committee would not be solely observers / minor players in existing change management projects which are mostly consultant-led, but would

instead be allowed to develop and run their own structure. They would also be given the administrative and operational support to allow them to develop as their own team.

b) The structure of the trainee committee would both parallel, and be integrated into, the School management structure.

c) The role of the trainee committee would not be 'contrived' but would be essential to the successful running of the School.

d) The trainees on the trainee committee would have a 'real' experience of the challenges of sitting between senior management and their peers, in a position of responsibility, where they have to make tough decisions.

e) The content and nature of the leadership development intervention would be tailored to the learning needs of the participants. This would be achieved by feeding information gained through the starting point focus groups (described in chapter 5) into the plans, and through constantly evaluating impact and making adjustments to the programme based on feedback from participants.

So, although also supported by seminars and group discussions, the main opportunities for developing leadership skills would come in the form of experiential learning (109) gained through participation in the activities described above. These principles were a vital base onto which the leadership development intervention was built, so if this programme were to be repeated or reproduced these would be important principles to adopt.

3.3.3 Application to the trainee committee

As described earlier in the chapter in sub-section 3.1.2, with the establishment of the London Specialty School of Paediatrics there was an opportunity to significantly expand the size, remit and influence of the trainee committee. A decision was made to use an application process, rather than nominations and election, in order to try to open up the process to trainees of all levels. An electronic advert was sent in September 2008, by preferred e-mail address, to all 992 trainees who were registered with the London Deanery as being current postgraduate trainees in paediatrics at that time. This advert included a ‘Questions and Answers’ document about the plans for the trainee committee (Appendix 3.1) and an application form. This form

asked applicants to complete some basic demographic details, give the name of a Consultant colleague (suggested, but not mandated, to be their educational supervisor) who was willing to support their application and to answer the following three questions:

1. What do you think you could add to the School Trainee Committee?
2. What are your biggest concerns about paediatric training in London and what are your ideas about how we could resolve them?
3. Are there any particular areas of work that you would like to get involved with? Why?

Applicants had 19 days to respond to the advert by submitting their completed application to a London School of Paediatrics administrative e-mail address. A reminder e-mail with the advert, information sheet and application form was sent out 7 days before the closing date.

Once all of the applications had been collated they were initially independently scored and shortlisted by the lead researcher (RK) and the Head of School (HC). Following this, all shortlisted applications were discussed in a face to face meeting of the lead researcher (RK), the Head of School (HC) and an experienced educationalist (FC) who was working on secondment to support the establishment of the School. These discussions included deciding on the placement of trainees into different sub-groups within the trainee committee. There was insufficient time and resource available to be able to interview shortlisted applicants as occurs with the majority of more formal job applications. The results of these processes are discussed in Section 3.4 below.

All successful applicants were e-mailed to inform them of the outcome and to invite them to take up a place on the trainee committee. All unsuccessful applicants were personally contacted, and invited to engage with the School by volunteering to take on a local 'Trust Rep' role to help with the communication of information between local paediatric departments and the School.

Once this process had been completed and acceptances had been received from all of the successful applicants the trainee committee could be formed. The plans for the trainee committee were finalised in December 2008. There were two London School of Paediatrics stakeholder meetings in January and February 2009 to actively 'launch' the School, and the trainee committee cohort first met in March 2009 and effectively became formally established at that point. The outcomes of this process are detailed in the results sub-section 3.4.1 below.

3.3.4 Participation in the trainee committee

Having established the trainee committee, the next phase was to provide the individuals on it with the leadership intervention that had been specifically designed for them. As introduced in sub-section 3.1.3, and following on from the discussions in chapters 1 and 2, the ethos of the leadership learning was that, rather than solely using the traditional model of taking participants out of work into a classroom setting, much of the learning would come from their participation in work-based projects and other activities within the London School of Paediatrics.

The trainees who accepted the invitation to join the inaugural trainee committee were given an introductory talk during their day-long induction event. This gave the opportunity to discuss and explore what was expected of them in the role, and to introduce the three core components of the leadership development intervention. It was agreed that the trainee committee members would do all that they could to attend all of the monthly centrally-held half-day sessions. However, it was recognised that for many of them, in acute or emergency focused posts, an attendance rate of 60-70% would be all that could be expected due to the need to participate in on-calls and night shifts. The trainees were also told about the fourth optional component of the leadership development opportunity – the leadership 360. The rationale behind this being optional is discussed below.

The expectation that they contribute to the discussions and decision making within the trainee committee meetings, and other School of Paediatrics meetings that they had the opportunity to attend, was also made. It was reiterated that 'representation of the views of colleagues' was an important role for all trainees on the committee. These experiences were designed to give the trainees who participated in the trainee committee the opportunity to gain competences in elements of the MLCF (see table 1.1) such as 'building and maintaining relationships', 'encouraging the contribution of others', 'making decisions', 'evaluating impact' and 'developing networks'. The trainees were each given a copy of the MLCF in order to help them to understand the context and terminology around each of the competencies.

3.3.5 Project-based leadership learning

In addition to being a member of the overall trainee committee each selected trainee was allocated to a sub-group based on their preferences indicated on the application form. This sub-group would give each trainee the opportunity to focus on specific areas of project activity, but also, with the much smaller numbers, would allow all of the trainees opportunities to experience leading projects. The different sub-groups were:

- Supporting trainees
- Curriculum delivery
- Assessment
- Communication & IT
- Recruitment & Workforce Planning
- Academic / Sub-specialty
- Faculty Development

The focus of the different projects within each sub-group was very varied, and covered a number of different elements of the Medical Leadership Competency Framework; examples of these are illustrated in more detail in the results sub-section 3.4.3. A number of the ideas for project-work emerged through discussions and reflections that took place during the focus-group sessions described in chapter 5. All of the projects were focused on improving the training experience for paediatric trainees in London, and hence aligning to the overall objectives of the Specialty School. There were 1½ hours allocated for each of the sub-groups to meet and discuss projects on a monthly basis, and each sub-group received informal support and mentoring from a consultant paediatrician acting in a Training Programme Director (TPD) role. Some trainees were able to involve themselves in 3 or more projects being run within each sub-group, others focused on a single project that ran throughout the year.

For many of the trainees, whose previous learning had solely been around the development of clinical competencies, this was their first experience of this type of work. The trainees were encouraged to write up their proposals in the style of a business case and, with the support of the sub-group Training Programme Director, to present them first to their sub-group and then to the School's Executive Board. This provided the opportunity for discussion, supportive challenge and feedback. The aim was that these opportunities would be representative of the sort of experiences, such as running improvement projects, redesigning their services or pitching for resources, that these trainees would be likely to encounter when working as consultants. Specifically, this work-based experience gave the trainees the opportunity to gain competencies in key MLCFs elements such as 'working within teams' and 'encouraging improvement and innovation'.

3.3.6 Workshop-based leadership learning

The first formal team-meeting of the London School of Paediatrics trainee committee in March 2009 was a day-long induction event that focused on introductions, scene-setting and

orientation. Two of the three focus groups described in section 5.3 were run during this event (see discussions in chapter 5 for the methodological implications, in particular around sequencing, of this). Following this induction, the School trainee committee was scheduled to meet for half a day in central London on a monthly basis, with three of these meetings each year superseded by a 'School Forum' event which was opened to all trainees and trainers across London. This meant that in the 12-month study period there were 9 trainee committee afternoons and 3 School Forum events, to which all trainee committee members were invited. Each of the trainee committee afternoons consisted of three components:

1. Agenda-led discussion across the full trainee committee about a wide range of training issues (90 minutes)
2. Small-group work within sub-groups to work on, and align, the project work described in sub-section 3.3.5 (90 minutes)
3. An interactive workshop on a key leadership concept. The aim was that these workshops had some alignment to one or more elements of the MLCF (75 minutes)

Some of the interactive workshops were facilitated in-house with others built around presentations from invited speakers. The first three delivered sessions that the trainees on the committee received were focused on understanding the importance of self-awareness and its interface with successful team-working. This aligns with the 'Demonstrating Personal Qualities' and 'Working with Others' domains of the MLCF. The content included an introduction to personality type through the use of Myers Briggs Type Indicator (MBTI) (121) assessments, work exploring Belbin's team roles (122) and participation in a team-based leadership simulation based around a management meeting scenario. The subsequent six workshops each used the insights of senior invited guests to help participants to develop a broader understanding of the wider healthcare system in England. Focusing on concepts around (1) health policy, (2) commissioning, (3) quality and safety, (4) rationing of healthcare, (5) understanding the context for change and (6) child public health they aimed to open minds and stimulate further thinking and exploration of ideas. These learning opportunities align with elements from the MLCF such as 'identifying the contexts for change', 'facilitating transformation' and 'managing resources' and gave the trainees the opportunity to address key competencies in these areas.

In summary the topics covered in these 75 minute long workshops were:

1. Self awareness, personality type and MBTI
2. Team roles and dynamics (including an introduction to Belbin's work)
3. Team-based practical leadership simulation with peer-observation and feedback

4. Introduction to the NHS and the wider healthcare system in England
5. Health policy and how it aims to lead to change
6. Commissioning – what it is, how it works, and what it hopes to achieve
7. Quality improvement and patient safety (highlighting Institute of Health Improvement)
8. Healthcare budgets, how money flows through the NHS and rationing
9. Public health and the preventative agenda; understanding the wider context for change

3.3.7 Designing a leadership 360 degree multi-source feedback tool

In addition to the learning from the workshops and the experiential learning gained from attending the trainee committee meetings and participating in innovative projects, the leadership development intervention had a fourth optional dimension. This was the opportunity to participate in a leadership 360 degree multi-source feedback tool that was specifically developed and built for this project.

Multi-source feedback (MSF) tools, which are sometimes called 360 degree feedback tools, are increasingly being used across healthcare to assess the performance of doctors and other healthcare professionals. There is a growing research literature around MSF, and this is well summarised in Wood et al's systematic review (123). In addition, the General Medical Council (GMC) has given multi-source feedback tools a prominent part in the process of appraisal. With appraisal emerging as the mainstay of the GMC's approach to the revalidation of doctors this is likely to widen considerably in the years ahead (124).

At the time of developing the project methodology there were no mainstream MSF tools focused on leadership development in use within the NHS. Although a number of commercial firms had developed leadership-based MSF tools they tended to be focused on senior leaders at board level, were often not completely relevant to healthcare and were prohibitively expensive to use. For these reasons, a decision was made to design and build a specific leadership-focused MSF tool that any of the trainee committee cohort could use on a voluntary basis to augment their leadership learning.

Wood et al's systematic review (123) was a helpful starting point for the design and structure of the 360, as based on their extensive review of the leadership literature, they proposed 10 key ideas that had come from the most successful published tools. These were to:

1. Develop a positive culture around the use of the tool

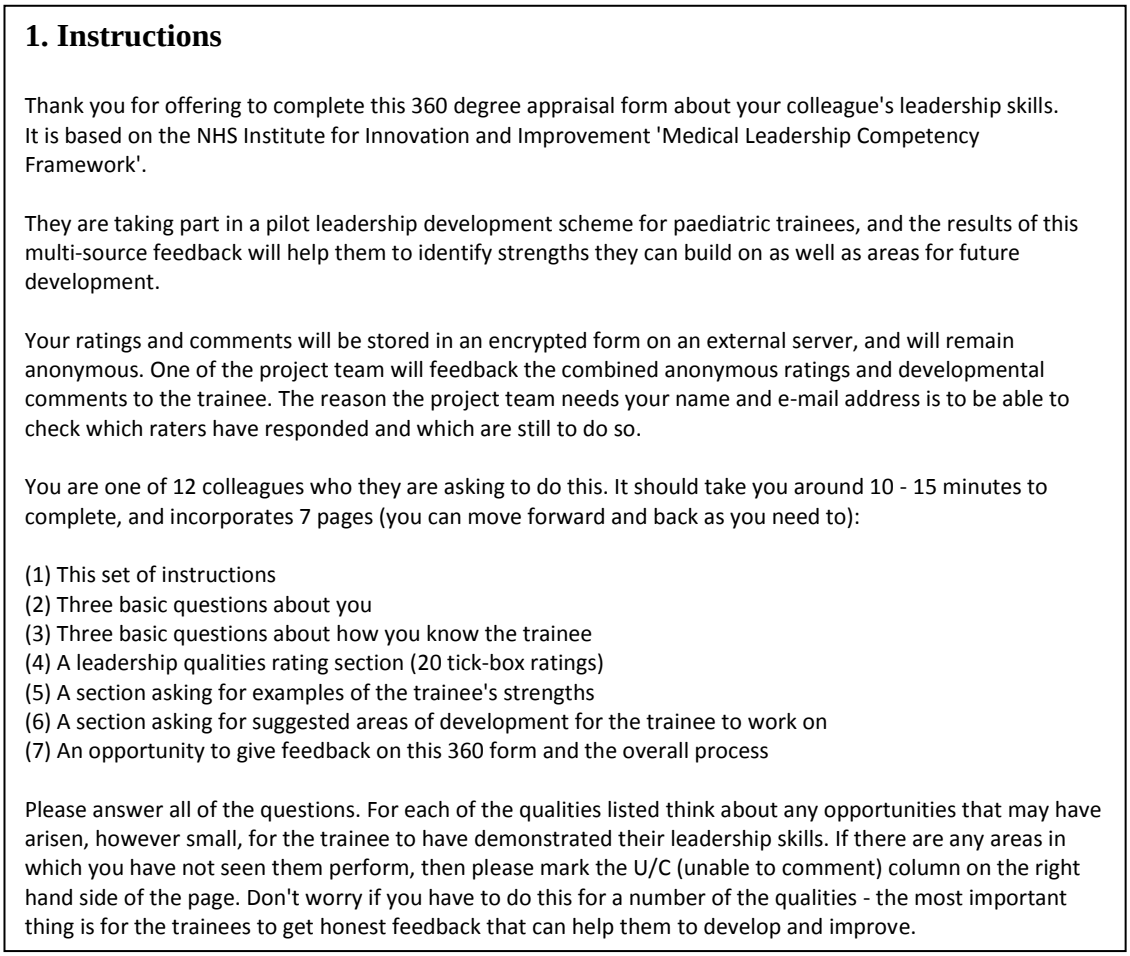
2. Be clear about the purpose of the tool
3. Clearly express any desired behaviours
4. Keep the number of items to be scored to a few
5. Keep the scale simple and fit-for-purpose
6. Use between six and ten raters
7. Compare the results with self-assessment
8. Train those giving feedback
9. Involve the assessees in the design and implementation
10. Use MSF to develop a shared understanding of the organisation

In addition to reviewing the leadership literature, the lead researcher (RK) had a number of informal conversations with senior leaders within the NHS in an attempt to learn from their experiences, and to input these into the design and planning. Non-healthcare avenues were also explored, the most valuable of which was a study-visit to meet the Head of Leadership at Diageo Ltd. Guidance was also sought from the team at the NHS Institute who developed the Medical Leadership Competency Framework. The emphasis on the importance of developing a tool that was simple, focused and fit-for-purpose was an important lesson learned from these conversations and visits. There was also the opportunity to build on the personal experience of using a broad, predominantly clinically focused, MSF tool called eSPRAT that was developed by the Royal College of Paediatrics and Child Health for use in all paediatric trainees.

The 360 degree multi-source feedback tool, which was named the 'Leadership 360' was designed by the lead researcher (RK), with the content validity of the tool achieved through basing the questions on the domains of the Medical Leadership Competency Framework (19). The questions were also influenced by some of the theoretical frameworks described in chapter 1, and in particular some of the ideas around followership and distributed leadership that came up through my reading on transformational leadership (22) and shared leadership (30). The questions in the Leadership 360 asked raters to report on the quality and frequency of different leadership behaviours (using a 5-point scale going from 'never' to 'always', and with an additional option not to comment) that they may or may not have observed in the individual whose Leadership 360 they had been sent. A total of ten questions were asked within a domain referring to 'Personal Qualities and Working with Others' and a further ten questions within a domain about 'Managing & Improving Services and Setting Direction'. There were also free text areas which asked raters to describe the trainee's leadership strengths and areas for development. The self-assessment form, which participants were asked to complete about themselves, followed the same structure.

The Leadership 360 was constructed using a professional account on the widely used 'Survey Monkey' web-based survey tool. This enabled sufficient functionality to be able to collate multiple responses for each individual, and also provided secure encryption of data. An information sheet for participants was developed (see Appendix 3.2) which was sent to the trainee committee with an open invitation to participate (see Appendix 3.3). It was made clear to the trainees on the committee that this was optional. The main rationale behind making this optional was that in any particular year around half of the trainees would be needing to complete a mandatory biannual workplace based multi-source feedback assessment that would also require them to ask colleagues to rate them. To have made this compulsory would have added to the burden of colleagues who were raters and may have led to unnecessary duplication.

Figure 3.2, which details the content on the introductory page of the Leadership 360, gives a useful summary of the ethos and processes behind it. The full version of the Leadership 360 (in printed, rather than web format) can be found in Appendix 3.4.



1. Instructions

Thank you for offering to complete this 360 degree appraisal form about your colleague's leadership skills. It is based on the NHS Institute for Innovation and Improvement 'Medical Leadership Competency Framework'.

They are taking part in a pilot leadership development scheme for paediatric trainees, and the results of this multi-source feedback will help them to identify strengths they can build on as well as areas for future development.

Your ratings and comments will be stored in an encrypted form on an external server, and will remain anonymous. One of the project team will feedback the combined anonymous ratings and developmental comments to the trainee. The reason the project team needs your name and e-mail address is to be able to check which raters have responded and which are still to do so.

You are one of 12 colleagues who they are asking to do this. It should take you around 10 - 15 minutes to complete, and incorporates 7 pages (you can move forward and back as you need to):

(1) This set of instructions
(2) Three basic questions about you
(3) Three basic questions about how you know the trainee
(4) A leadership qualities rating section (20 tick-box ratings)
(5) A section asking for examples of the trainee's strengths
(6) A section asking for suggested areas of development for the trainee to work on
(7) An opportunity to give feedback on this 360 form and the overall process

Please answer all of the questions. For each of the qualities listed think about any opportunities that may have arisen, however small, for the trainee to have demonstrated their leadership skills. If there are any areas in which you have not seen them perform, then please mark the U/C (unable to comment) column on the right hand side of the page. Don't worry if you have to do this for a number of the qualities - the most important thing is for the trainees to get honest feedback that can help them to develop and improve.

Figure 3.2 Introductory Page of 'Leadership 360' multi-source feedback tool

Members of the trainee committee cohort who opted to participate in the Leadership 360 were given a weblink (URL), which they were asked to send out to 12 self-nominated colleagues with whom they work. For comparison they were also asked to use a second web link to complete a self-assessment version of the tool. Both were set-up so that it was not possible for any single participant to give more than one set of results.

Once a 4-week collection period had expired and a minimum of 8 responses had been received - the recommended level from an important national study (125) – participants were encouraged to use the results as part of a developmental discussion with a senior colleague. It was suggested that this might be their educational supervisor or a mentor, but the option was also given to arrange an appointment with the lead researcher (RK) or the Head of School (HC) to discuss results. The aim of this reflective discussion was to support their leadership learning through widening insights and opening up new perspectives. As with the overall participation in the Leadership 360, this meeting was encouraged but optional. This opportunity to participate in the Leadership 360 covered many of the leadership competencies within the first domain of the MLCF, 'demonstrating personal qualities', with particular focus on developing self-awareness.

3.3.8 Designing the evaluation strategy

This section gives an overview of the evaluation strategy designed to explore the impact of the leadership development intervention. With this project involving only a small cohort of participants, and run over a relatively short period of time (12 months in the phase covered by this thesis), the evaluation methodology needed to be carefully designed to be in with a chance of picking up any real impact of the leadership intervention.

3.3.8.1 *Previous experiences of evaluating educational interventions*

Prior to undertaking this thesis, the framework that I had previously used to guide the evaluation planning of educational interventions I was developing was Kirkpatrick's Learning Evaluation Model (89), which was originally described by Donald Kirkpatrick in 1975. This evaluation framework works at four levels: reaction, learning, behaviour and results. Table 3.1 provides a summary of the model.

Level	What is measured	Description	Evaluation methods
1	Reaction	How the participants felt about their learning experience	- Assess the level of participation in the different elements of programme - Identify the challenges that emerged from the programme - Ask which elements of the programme worked well and which you would change or do differently
2	Learning	The measurement of the increase in knowledge during the programme	- Assess the knowledge and skills learned during the programme
3	Behaviour	The extent to which the learning applied to the workplace	- Identify attitudes and behaviours towards other professional groups and towards wider organisational issues before and after participation in the programme - Evaluate how participation in the programme has changed these behaviours
4	Results	The effect on the organisation or environment as a result of the learning experience	- Identify whether specific project outcomes were achieved - Identify the impact on department, colleagues and organisation

Table 3.1 Kirkpatrick's four levels of training evaluation, adapted from Phillips (1996).

Whilst levels 1 and 2 are usually reasonably straightforward to achieve through traditional evaluation methods involving the individual participants, exploring the extent to which a programme has had impact at levels 3 and 4 is much more challenging. Looking at this project through the 'Kirkpatrick lens', the quantitative aspects of the evaluation could be said to be focused on an assessment of knowledge at level 2, whilst the qualitative work to explore participants' reactions was seeking to evaluate the impact of the programme on attitudes and behaviours (level 3). There might also be the hope that any perceived impact at organisational level (level 4) would be picked up through the focus groups, but a more comprehensive evaluation of the wider impact was beyond the remit of this thesis.

Despite its very widespread use as a framework for evaluation of training initiatives, there are a number of limitations to this model. Yardley and Dornan argue that Kirkpatrick's levels, which they highlight were introduced to evaluate training in industry, involve so many implicit assumptions that they are suitable for use only in relatively simple instructional designs, short-term endpoints and for beneficiaries other than learners (126). It can also be argued that this four-level model presents an oversimplified view that does not consider individual or contextual influences that may come into play. It can be criticised for over-focusing on training as an event, where instead learning is perhaps better thought of as a process that is connected, collaborative and continuous (127). The model also assumes that there is a causal linkage between the four levels inferring that a better student reaction leads to greater learning and thus more organisational impact. Two meta-analyses of training evaluation studies using Kirkpatrick's framework (128)(129) have found little evidence that this is the case. Finally, there is an assumption that the information gained from within the higher levels of the model is more important that the information from Levels 1 or 2. However, it can be argued that the weak conceptual linkages inherent in the model do not provide an adequate basis for this assumption (130).

There are a number of other models that have emerged since Kirkpatrick's model that can be used to evaluation educational interventions. Many of these move towards evaluating against the intervention's outcomes and processes (131)(132). Kaufman, Keller and Watkins (133) reframed Kirkpatrick's four levels and also included a 5th dimension, societal contributions. They also make a shift towards evaluating the organisational factors that underpin learner reaction, experience and learning, and an attempt to better connect performance to expectations. However, in many ways the criticism about the causal linkages described with reference to Kirkpatrick apply here as well. Other researchers have moved towards an approach that builds a logic model around the educational initiative (134)(135). This aims to capture emergent outcomes and to then investigate why and how the intervention actually gets

to these outcomes. This can then inform the development and evaluation of future versions of the initiative. One of the most commonly cited criticisms is that, if oversimplified, the logic model can underplay the nonlinear complexity of most educational contexts (136). The decision making around the design of the research methodology in this study is described in detail below.

3.3.8.2 *Evaluation or research?*

Reflecting on my previous use of Kirkpatrick's model, the question as to the differences between evaluation and research arises. This question continues to be widely debated and there are many diverse and at times conflicting perspectives (137) (138).

There are those who argue that evaluation and research are fundamentally different. They cite the importance of 'valuing' to evaluation - this is the focused application of evaluation findings (ie local and specific, rather than being more generalizable). They also highlight the important role of theory in research compared with a much lesser role within evaluation. In effect evaluation determines the merit, worth, or value of things. It has been argued that the evaluation process identifies relevant values or standards that apply to what is being evaluated, performs empirical investigation, often utilising different techniques from the social sciences, and then integrates conclusions with the standards. The result is an evaluation plan (139). Contrasting this, research can be defined as the search for knowledge, or as any systematic investigation, to establish facts, solve problems, prove new ideas and develop new theories. Research usually follows a scientific method.

The alternative perspective comes from those who view evaluation as a *form* of research, where research methods are used to answer practical focused questions in a timely fashion. They argue that the commonly cited differences between evaluation and research only apply to some types of research. There is suggestion that one might want to think of evaluation as an applied social science (140). Patton describes evaluation as the systematic collection of information about the activities, or outcomes, of programs in order to make judgements about the program, improve program effectiveness and/or inform decisions about future programming (141).

The open nature of the thesis title 'Clinical leadership: can the skills be learned by trainee paediatricians?' is perhaps suggestive of this work being research, and as described in the sub-section that follows, this question can be repositioned as a research hypothesis. However, it

can also be argued that this thesis fits well into Patton's definition – it describes the evaluation of a specific leadership development programme, to determine its value. On balance this work is probably best described as an evaluation, but one that is underpinned by a thoughtful focus on research methodology. Key to this is developing a methodological approach that evaluates against the project's objectives.

3.3.8.3 *The rationale for a mixed methods approach*

At this stage it is worth reflecting back on the overall aim of this project - namely to explore whether paediatricians in training can develop leadership skills through participating in a specifically designed leadership development initiative. Within this aim, it could be argued that there is a hypothesis, or maybe a theory, that suggests that "paediatricians in training can develop leadership skills through participating in a leadership development initiative". The research question effectively needs to test, or explore, this hypothesis. With the nature of this question, the subjects involved, and the type of intervention developed, a research approach that combines aspects from both qualitative and quantitative methodologies feels likely to be the most appropriate direction to take.

Lingard et al's introduction to grounded theory, mixed methods and action research was a helpful place to start in considering the potential role of each of these different research modalities. This paper describes mixed methods research as:

> '…combining elements from both qualitative and quantitative paradigms to produce convergent findings in the context of complex research questions. There are tensions between these methods in terms of their values and processes, but these very tensions can generate new insights. (142)'

At this explorative stage grounded theory (143) was also considered as a potential methodology for this study. As explained in the Lingard paper, grounded theory takes a different approach, in that it has an iterative design that involves cycles of simultaneous data collection and analysis. These initial sequences of data analysis lead on to further adaption of the data collection proceed. Indeed, the sample used in this design is not set at the beginning, but is selected as the process unfolds. Participants are chosen for their ability to confirm or challenge the theory as it emerges. In many ways, compared to traditional research methods, this is like a hypothesis that has been engineered in reverse. This method of grounded theory was first described by two sociologists, Glaser and Strauss in 1967 (143), and emerged at a time when qualitative research was generally considered to be 'unscientific'. Grounded theory is perhaps best defined as a systematic approach to generating theories or hypotheses from data, using thinking and questioning techniques that are both deductive and inductive. It is

important to recognise that this is different from narrative or ethnographic approaches, which are also increasingly recognised methods of qualitative research. How formally and rigorously the mechanisms and methods of grounded theory are applied to data is a subject of significant academic debate, which even led to a significant divergence of views between Glaser and Strauss themselves. Whilst Glaser continued to promote an approach that was interpretive, set in the wider context and self-developing (144), Strauss proposed an even more thorough, systematic analysis (145). In his method every piece of data would be coded, all possible meanings for each piece of data would be considered and only then would a conditional matrix be constructed, from which a theory could be described.

Lingard et al argue that grounded theory is appropriate when the study of social interactions or experiences aims to explain a process (142). They suggest that it is not appropriate as a methodology to test or verify a theory. For this reason it can be argued that it is *not* the right methodology for the research within this thesis.

3.3.8.4 *Introduction to experimental design*

Having explored the case for a mixed methods approach the next stage was to determine the exact experimental design for this thesis. My learning and understanding of the different methodological designs that could be used came from Campbell and Stanley's Experimental and Quasi-Experimental Designs for Research (146), Creswell's text on research design (147) and Greene et al's conceptual framework for mixed-method evaluation designs (148).

The starting point of exploration in this area was to consider the value and limitations of a cohort (in this case the trainee committee) having some sort of pre-intervention test and then also a post-intervention test. Campbell and Stanley call this 'Design 2 – the One-Group pre-test - post-test design' (146). While this design is very widely used in educational research, and can have some value where there are not alternatives, it is described as a 'bad example' in Campbell and Stanley's original paper. This is due to the significant number of confounding variables that can jeopardize the internal validity of the results. These include potential concerns with:

a) History – other change over time producing events in addition to the intervention
b) Maturation – changes in the cohort group with the passage of time, independent of the impact of the intervention
c) Testing – the effect of the pre-test itself. In this project it can be argued that the pre-test self-assessment and focus group participation are in effect part of the intervention.

d) Instrumentation – where there are autonomous changes in the measuring instrument or tool between the pre-intervention test and the post-intervention test

This design also creates issues with external validity due to what could be described as 'interaction effects'. These are interactions between the intervention and the selection process used to select the cohort group, and between the intervention and the testing process. These factors that are jeopardizing external validity lead to difficulties with generalisation.

As introduced in chapter 1, and described in more detail in chapters 4-7, this study looked to use a comparator group to strengthen the research methodology. Campbell and Stanley detail a number of true experimental designs where the use of a comparator, or control group, reduces the issues around internal sources of invalidity, but also what they describe as 'quasi-experimental designs'. These are situations where the researcher has control over when the data collection procedures take place, but much less control over the timing and nature of exposure to the experimental stimuli itself. These quasi-experimental designs are deemed worthy of use where better designs are not possible or feasible.

In considering the potential use of a control group, the pragmatic difficulties of whether a true control group would be able to be randomly selected, and then put through pre-test and / or post-test testing became a dominant factor. Although this approach weakens the research design, and can thus limit the conclusions made, provided due considerations are made to the recognised jeopardising factors, it is still a legitimate approach to take. The main difficulty in this research was the practical difficulty of finding a true control group, and then putting them through a post-test alongside the trainee committee cohort. For this reason a comparator group approach was chosen.

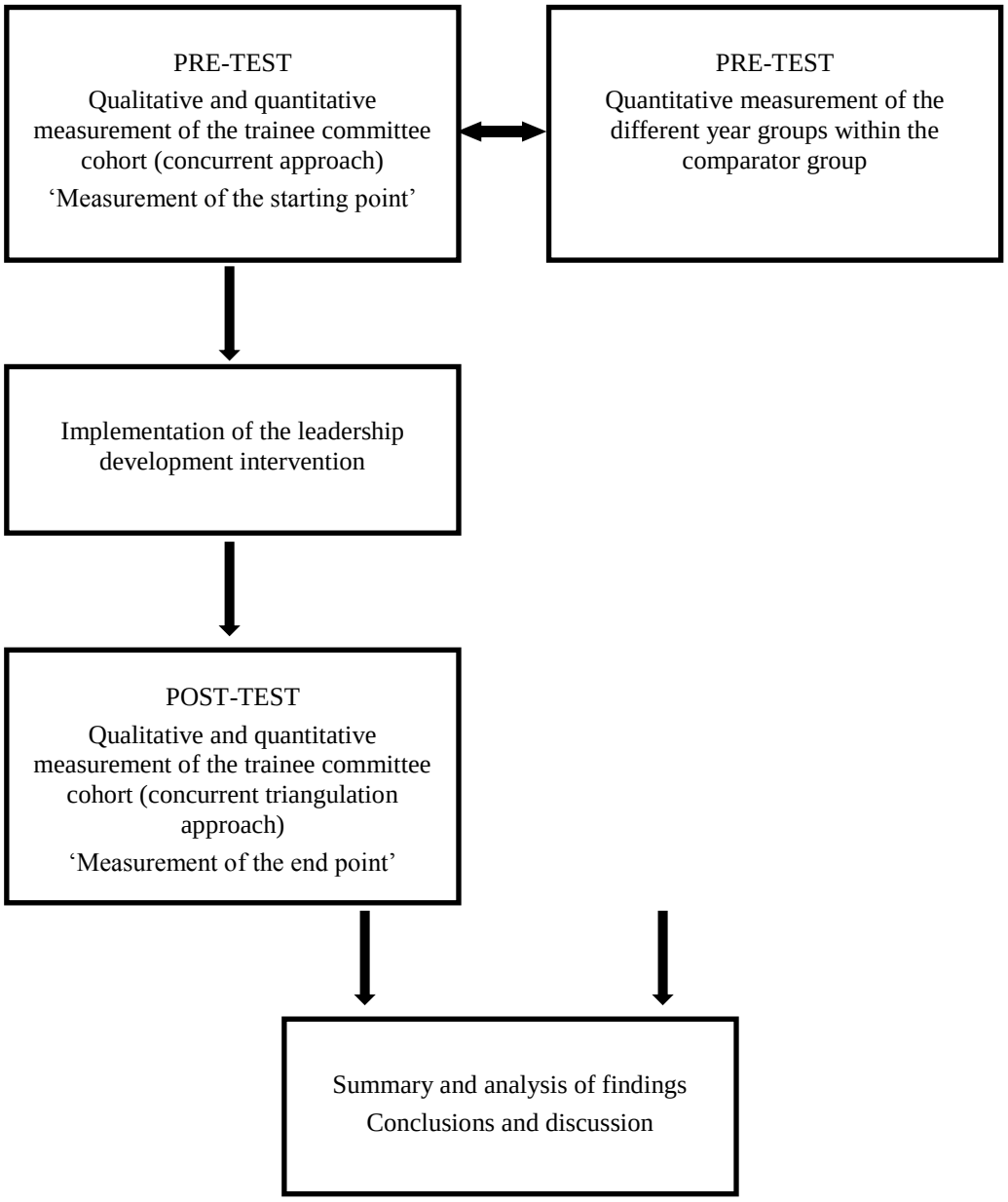
In their main chapter on quasi-experimental designs, Campbell and Stanley state:

> "it is in the spirit of this chapter to encourage 'patched-up' designs, in which features are added to control specific factors, more or less one at a time (in contrast with the neater 'true' experiments, in which a single control group controls for all of the threats to internal validity.(146) "

Within this chapter, their Design 12 - the separate sample pre-test - post-test design and Design 13 - the separate sample pre-test - post-test Control Group Design both have features that helped to influence the final design of this study, and interpret the results within the context of the methodology used. In effect the study became a one-group pre-test – post-test design, but also utilised a comparator group. This comparator group consisted of trainees from a range of year groups, who were also tested at the time of the trainee committee cohort's pre-test. The

aim of this was to eliminate or reduce some of the four main jeopardising factors that threaten internal validity listed on the page above. The detail of this is discussed in Chapters 4 to 7.

Figure 3.3 Flowchart summarising the overall study design:



So, in summary, three research methodologies were designed to answer the main research question:

1. Systematic review of the literature for medical leadership development interventions – this search for empirical evidence was described in detail in chapter 2, and was designed to give the project context, as well as informing the design of the intervention in this project. This systematic review was also contextualised in the wider leadership literature outlined in chapter 1. This approach allows for comparison of the research findings from this project with learning and research from across the international literature, and thus helps with generating meaningful conclusions from the research results.

2. Quantitative measurements using the self-assessment leadership capability surveys – these are described in detail in chapters 4 and 6. The aim of this part of the evaluation strategy was to provide the pre-test 'before' and post-test 'after' measurements of leadership capability of the trainee committee cohort, and also to compare their baseline pre-test assessment scores with those of a wider cohort of peers; the comparator group. Unfortunately it was not possible to do a post-test assessment of the comparator group.

3. Qualitative analyses of the trainee committee cohort's perceptions prior to, and after, the implementation of the leadership development initiative – these are described in chapters 5 and 7. The aim of this qualitative analysis, with data gained through focus groups, was to learn more about the attitudes, perceptions, feelings and behaviours of the participants. Once themes emerged this information could then be used to gain a deeper understanding of the 'pre-test' and 'post-test' quantitative data.

A further important discussion point with regards to the study design is the sequencing of the data collection and data analysis, and the different methods that can be used to integrate data. These issues are also addressed in chapters 5, 7 and 8. There is an extensive literature describing different strategies and approaches to mixed-methods research (147)(149)(148)(64) involving the timing, weighting, mixing and theorizing of the qualitative and quantitative data collection. These four factors help to shape the procedural aspects of a mixed methods study, and lead to a number of different possible designs. In this study the concurrent triangulation approach, which is probably the most familiar mixed methods approach, has been adopted. The aim was to collect data concurrently so that the two databases could be compared to see if there is convergence, differences or some combination of the two. Some of the practicalities of

this are discussed in chapter 5. Although this approach can lead to well-validated and substantiated findings, Creswell notes that there are limitations to this approach (147). He argues that it requires significant effort and expertise to study a single phenomenon with two separate methods, and it can be difficult to compare the results of the two analyses using data of different formats. It can also be difficult to work out how to deal with discrepancies in the data collected by the two methods.

Finally, there are different possibilities in terms of how mixed methods data is integrated. The word 'triangulate' is often used with range of meanings, far wider than its original definition. Greene et al (148) suggest a number of different methods through which data can be integrated. The approach taken in this study is discussed in the subsequent chapters.

3.3.8.5 *Consideration of an action research approach*

One other possibility that was considered, and ultimately rejected, was whether or not an action research approach could have been taken within this thesis. Action research, which is sometimes also called community based research, participatory action research or collaborative enquiry, is research that is designed, implemented and run by the participants in partnership with the researcher. It is an iterative approach where the participants and researchers act together within the context of a problem to explore and implement change. The work and research is set within a set of mutually agreed ground rules (142). Although the involvement of the lead researcher (RK) in the trainee committee was fundamental to its successful launch and ongoing development, his role within the trainee committee and that of the participants was sufficiently different that it would have been difficult to argue that the research agenda and activities were shared by researchers and partners. It could also be argued that the content of the intervention was not necessarily as strong an issue as the sorts of community-based, collaborative action research projects that have had success within healthcare (150).

3.4 Results

Whilst chapters 4 to 7 each describe a different section of exploratory research, the story of which fairly naturally fits into the traditional 'methods, results, conclusions and discussion' framework, this chapter has brought together a number of related strands to describe the establishment of the trainee committee and the development of the leadership intervention. Therefore this section has been used to describe results and outcomes of the applications to the trainee committee, the participation in the trainee committee and uptake of project work, the participation in workshops and the uptake in the Leadership 360.

3.4.1 Applications to the trainee committee

This first sub-section describes the results of the application process to the trainee committee. 992 trainees, who were registered with the London Deanery as being current postgraduate trainees in paediatrics at that time, were sent the advert detailing how to apply to the trainee committee. 65 completed applications were received prior to the deadline, along with 2 forms that were incomplete and which were therefore rejected at this stage. It had been expected that around 20-25 applications would be received, so this response was three times more than expected. As a result of the strength of the applications and energy they exuded, the planned structure of the trainee committee was adapted at this stage so that it could increase its membership from 20 to 32. The initial shortlisting process reduced the number of potential trainees down from 65 to 52. These applications were then reviewed, using the processes detailed in sub-section 3.3.1 above, and 32 successful applicants were chosen to be members of the trainee committee. All 32 trainees who were invited to become members of the inaugural trainee committee accepted the offer. Of note there was a strong distribution of trainees across the different levels of training, which was one of the main aims of choosing an application form as the selection methodology. These levels of training, the spread across different hospitals / Trusts and the allocation of sub-groups are summarised in Table 3.2. Each trainee was given a unique 'participant code' in order to maintain their anonymity.

Work Area	Participant Code	Level	Hospital / Trust
Supporting Trainees	1	ST1 (N)	Hillingdon
Curriculum Delivery	2	ST1 (N)	Whipps Cross
Communications + IT	3	ST1 (S)	Mayday
Selection, Recruitment + W Planning	4	ST1 (S)	Lewisham
Curriculum Delivery	5	ST2 (KSS)	Royal Surrey
Assessment	6	ST2 (N)	Royal Brompton
Academic / Research / Sub-spec	7	ST2 (N)	Barts + The London
Communications + IT	8	ST2 (S)	Kings
Assessment	9	ST2 (S)	Guys + St Thomas
Supporting Trainees	10	ST2 (S)	Guys + St Thomas
Communications + IT	11	ST3 (N)	Whittington
Supporting Trainees	12	ST3 (N)	Barts + The London
Academic / Research / Sub-spec	13	ST3 (N)	Barts / ICH
Communications + IT	14	ST4 (N)	Homerton
Assessment	15	ST4 (N)	Ealing
Curriculum Delivery	16	ST4 (N)	NWP
Curriculum Delivery	17	ST4 (N)	Whipps Cross
Selection, Recruitment + W Planning	18	ST4 (S)	Farnborough, Kent
Selection, Recruitment + W Planning	19	ST5 (N)	Barnet
Selection, Recruitment + W Planning	20	ST5 (N)	Whittington
Curriculum Delivery	21	ST5 (S)	Kingston
Supporting Trainees	22	SpR Y2 (N)	North Middlesex
Curriculum Delivery	23	SpR Y2 (S)	Ashford
Faculty Development	24	SpR Y2 (S)	East Kent

Assessment	25	SpR Y3 (N)	GOSH
Supporting Trainees	26	SpR Y3 (N)	St Mary's
Academic / Research / Sub-spec	27	SpR Y3 (N)	Barts + The London
Faculty Development	28	SpR Y3 (N)	Chelsea + West
Audit + Research Trainee Lead	29	SpR Y3 (N)	Chelsea + West
Supporting Trainees	30	SpR Y3 (S)	St George's
Assessment	31	SpR Y4 (N)	UCL / UCLH
Academic / Research / Sub-spec	32	SpR Y4 (S)	GOSH

Table 3.2 Categorisation of successful applicants to the trainee committee by level of training, hospital where they were based and sub-group placement

As described in the introduction to this chapter, a key principle of the trainee committee was that the ideas and vision of the committee were integral to the strategic direction of the School. In order to follow this principle from the very outset, a thematic analysis of the free text answers to the 'What are your biggest concerns about paediatric training in London and what are your ideas about how we could resolve them?' questions within the trainee committee application forms was undertaken. This work, which was led by the two educational fellows (AR and RH) with support from the lead researcher (RK), the Head of School (HC) and an educationalist (FC) is not directly related to this research and so has been summarised in Appendix 3.5 rather than being presented within the main chapters of the thesis. The analysis was presented to the School Executive Committee and back to the Trainee Committee to be used as evidence of areas of opportunity and concern that could be focused on.

3.4.2 Participation in the trainee committee

Each of the 32 trainee committee members participated in the trainee committee to varying extents. Formal attendance lists were not maintained contemporaneously, but on retrospectively reviewing of the e-mail apology lists it can be estimated that attendance ranged from 3/12 (25%) to 10/12 (83%) of sessions. The majority of trainees on the committee were able to attend 7 or 8 (58% – 67%) of the sessions in total. There was no formal measure of the specific level of participation of each individual, but all 32 trainees on the committee had the opportunity to discuss training issues and represent their colleagues with the senior training leads (Training Programme Directors) and the Head of School.

3.4.3 Project work uptake

All 32 trainee committee members participated in one or more projects or workstreams within the different sub-groups. The origin of the project work was very varied. Some work was passed downwards from the Head of School or School Board, but the majority was generated

from the energy and imagination of the trainees involved. In either situation the projects were trainee led, with senior advice and guidance available from a consultant level Training Programme Director. Their role was around enabling and facilitating the trainees to make progress with their work and to get the best possible learning from those experiences.

Each of the projects had an emphasis on improving the quality of training within paediatrics in London, but in all cases was contextualised by the acknowledgement of the relationship between high quality training and safer, more efficient patient care. Trainees were encouraged to innovate and take risks with projects, but to have a strong evaluation strategy in order to remain focused on project outcomes, and to be able to learn from any mistakes. Project work was undertaken flexibly, with the majority of the work done in the trainees' own time. As a basic illustration of the areas explored, some examples of initial projects that were undertaken include the:

- Establishment of a web-based curriculum-linked resource tool
- Design and implementation of trainee led training guide
- Improvement of recruitment processes and careers guidance to increase the number and quality of applicants to paediatric training
- Development of systems for more collaborative trainee-led research and audit between Trusts
- Redesign of the processes involved in application for 'Out of Programme Experience'
- Development of a curriculum delivery strategy for the School

The aim was that the experience of participating in and leading these projects would provide the trainees with a very pragmatic, experiential form of leadership learning, linked to a number of the different elements of the Medical Leadership Competency Framework (19). The extent to which this was successful or not is explored in chapters 6 and 7.

3.4.4 Workshop uptake

Despite the potential learning from the workshop topics outlined in sub-section 3.3.6, it is important to re-emphasise that, as described in the introduction to this section, and the early sub-sections within the chapter, it was the work-based experience and project work that were envisaged to be the mainstay of the leadership development intervention. In addition to the evidence from the MLCF, and the learning from some of the papers found in the literature review (100)(111), one very significant reason for this was a strong realisation of the

difficulties that the members of the trainee committee had in attending the trainee committee meetings on each occasion they were held. With the exception of two trainees who were on academic placements, all other trainees on the committee were part of on-call rotas within their Trusts. This meant that when these trainees were on-call, or preparing to do the night shift, they were not able to attend the trainee committee afternoons.

Formal attendance lists were not maintained contemporaneously, but on retrospectively reviewing the e-mail apology lists it can be estimated that attendance at the 9 workshops ranged from 3/9 (33%) to 9/9 (100%) of the sessions. The majority of trainees on the committee were able to attend between 5 and 7 (56% – 78%) of the workshops in total. Summary notes were made from each session so that trainees who were not able to attend could review the key points and potentially gain deeper learning through exploration of accompanying reading lists, or through discussions with their peers.

As with the sub-section above relating to project work, the extent to which these workshops were of value or not is explored in chapters 6 and 7.

3.4.5 Leadership 360 degree multi-source feedback tool uptake

As highlighted in sub-section 3.3.7, participation in the Leadership 360 was entirely voluntary. In total the opportunity to trial the Leadership 360 was taken up by 17 of the 32 trainees who were selected onto the trainee committee. This equates to a 53% uptake rate. Many of the trainees also had to undertake a compulsory generic multi-source feedback assessment, called eSPRAT, as part of the annual assessment process for their training programme. All of the trainees who declined the opportunity to participate in the Leadership 360 cited this as the reason why they did not participate. Their main concern was that they would not get enough rater feedback across the two surveys if they took them both on.

Of the 17 trainees who were sent individualised Leadership 360 web-links, 11 received between 10 and 12 responses, 4 received between 6 and 8 responses and 2 trainees only received 2 and 4 responses respectively. Once completed trainees were sent an electronic summary of all of the responses received from their colleagues. This was available for them to read and reflect on, but in addition the majority of the trainees incorporated discussions about their feedback into supervisions with their educational supervisor. Four trainees arranged for feedback meetings with the lead researcher (RK) and these were held through face-to-face discussions in the weeks that followed the completion of the collection of responses.

As the use of multi-source feedback tools within medicine continues to expand there is a growing medical education literature exploring the reliability, validity and feasibility of their use as workplace-based assessments (151)(152)(153)(154)(155)(156). Overall the use of multi-source feedback tools employing medical colleagues, other healthcare colleagues, and patients as a method to assess doctors in practice can be argued to show high reliability, validity, and feasibility. There are arguments about the merits of self-nominating raters or having them selected externally. In this study, with such small numbers of participants, formal critique and review around the validity, reliability and generalizability of the Leadership 360 was not undertaken. As an optional resource for the individual trainees within the trainee committee cohort to utilise, it was important to describe the processes undertaken with the Leadership 360, but it was never intended to be a core component of the leadership development intervention. As with the sub-sections above relating to project work and workshop attendance, the question as to whether or not participation in the Leadership 360 was of value to the 17 trainees who enrolled on it is explored in chapters 6 and 7.

3.5 Conclusions

This chapter has explored the processes involved in the planning and implementation of an innovative leadership development initiative, which was specifically designed to support the learning and development of the newly established London School of Paediatrics trainee committee. This leadership development intervention consisted of:

(1) involvement in the work of the trainee committee
(2) project-based work run through the sub-groups
(3) attendance at specifically designed leadership workshops
(4) the opportunity to participate in a Leadership 360

Trainees were not able to participate in every aspect of each of these opportunities due to other work commitments, although the programme was left as flexible as possible to accommodate this reality of a work-based programme. A mixed methods methodology was designed to evaluate the impact of this leadership development initiative on the trainees who participated in it. This has been outlined in sub-section 3.3.8 above and is illustrated in detail in chapters 6 and 7.

3.6 Discussion

The essence of this project was to design and implement a leadership development programme that, underpinned by the concepts within the Medical Leadership Competency Framework, supported the leadership learning of a cohort of paediatric trainees who had been selected to

form the inaugural London School of Paediatrics trainee committee. As described in section 3.1 these developments were initiated at a time of significant change within postgraduate medical training. Although this change generated some turbulence in the system, creating new structures and processes, and a significant degree of uncertainty, it also provided an opportunity for new initiatives. The establishment of the trainee committee and the associated leadership development programme came out of this opportunity.

The selection of the trainee committee was carried out through well managed processes that are detailed in sub-section 3.3.3, but due to the surprisingly large number of applicants some adjustments had to be made to the structure of the trainee committee in order to gain the most from this opportunity. The scoring and the shortlisting processes aimed to explore commitment to the programme alongside rewarding trainees who had a strong track record of designing and implementing projects. As with any selection process, aspects of this were imperfect and if run again, with additional resource available, it may be worth relooking at this process to incorporate interviews as an additional selection step. Above all, an interview might be a helpful way of exploring commitment to the programme, discussing specific individual learning needs, and orientating the trainee applicant to the ethos of the role. A second dimension was the role of 'application' to the trainee committee rather than 'election' which is the traditional format taken, for example, by many of the Medical Royal Colleges. As described in sub-section 3.1.3 the aim of this approach was to broaden the representation to include a strong balance of dynamic junior trainees who were only just starting out in their programmes. To this end, as illustrated in Table 3.2, this was extremely successful.

As highlighted in chapter 1, most trainees in the UK have had little or no leadership development training. Where individuals have had opportunities these have tended to be classroom based and focused on the more senior trainees. Having successfully recruited a very dynamic and heterogeneous cohort of 32 trainees to form the inaugural trainee committee, it was important to design and implement a leadership development intervention that would provide opportunities for leadership learning for them all. The aim was to develop an initiative that reflected some of the thinking from the wider leadership literature and the medical leadership systematic review (100)(94)(102)(106). Particularly important was to build this initiative around the principles described by Pearce, Conger and other authors writing about shared leadership (30)(117)(31). Based on the rationale described in sub-section 3.1.3 this was work-based and also highly flexible and pragmatic in acknowledging that there would never be occasions when all 32 individuals could come together. The intervention was also very low-cost and therefore, if shown to be of value, would have a good chance of long-term

sustainability and spread. Early opportunities for this would be spread to other paediatric Specialty Schools across the UK, or to non-paediatric Specialty Schools within London.

The main difficulty of this approach was from a research perspective, where the 'blurring around the edges' of which experiences were inside the defined leadership development intervention and which were not was complex. The risks with this were three-fold. Firstly, there was the potential that any improvements or change following the leadership development intervention may not be clearly picked out by the research methodologies chosen. Secondly, any changes picked up by the evaluation may actually be attributable to some other factor outside the influence of the leadership development intervention; without a control group this may be difficult to detect. Thirdly, although the intervention had a degree of flexibility and bespoke approach for trainee committee cohort, when thinking about sustainability and spread, reproducing exactly the same experience for other cohorts would be extremely difficult. These issues, hazards and jeopardizing factors are explored in chapters 4, 6 and 8.

With regards to the optional Leadership 360 it was noted at the time of developing the project methodology that there were no mainstream MSF tools focused on leadership development in use within the NHS. Since developing this tool this situation has changed, and in addition to the marketing of further commercial leadership tools, two different healthcare leadership resources have been developed and made available to NHS staff. The first of these, the NHS Leadership Framework self-assessment tool (157) (initially launched in 2011), is not actually a multi-source feedback tool, as it only focuses on self-assessment. However, in a similar way to the self-assessment component of the Leadership 360 designed for use in this project, the tool is based on the domains and elements of the NHS Leadership Framework (158), the difference being that it is pitched at all healthcare professionals and not just doctors. The second resource was developed on behalf of the NHS Leadership Academy by the talent and career management company Right Management. This resource is a multi-source feedback tool built around the content of the NHS Leadership Framework and designed to provide a practical way of assessing leadership behaviours. Unlike the self-assessment tool there is a reasonably significant charge to using this feedback tool.

It remains to be seen whether there will be wholesale uptake of leadership-focused multi-source feedback tools such as the Leadership 360 designed and used in this project. In this project, as noted in sub-section 3.4.5 above, uptake of this optional tool was only 53% and in effect only just over one third of the cohort group had enough respondents for it to be worthwhile. Within chapter 7 there is some trainee-led discussion of this issue during the post-

intervention focus groups. The concern for trainees is that in most specialities they are already mandated to undertake annual multi-source feedback assessments and there are significant issues in some areas of trainees struggling to get enough responses. The more common and widespread their use becomes there is a risk of 'rater fatigue' and that insufficient people respond, and some of those that do spend less time on their feedback than perhaps they previously did.

Sub-section 3.3.8 describes in detail the thinking and considerations made in designing the overall evaluation strategy for the project. It highlights the conclusion that, on balance, this work is probably best described as an evaluation rather than research, but one that is underpinned by learning from research methodologies and techniques. This sub-section also describes the thinking and rationale behind the evaluation design, which is illustrated in Figure 3.3. The use of a comparator group, who only undertook the pre-test, to attempt to strengthen the research methodology is not without controversy. As highlighted in sub-section 3.3.8.3, Campbell and Stanley (146) describe a number of true experimental designs where the use of a comparator, or control group, reduces the issues around internal sources of invalidity, but they also illustrate what they call 'quasi-experimental designs'. These designs are utilised when the researcher has control over when the data collection procedures take place, but has much less control over the timing and nature of exposure to the experimental stimuli itself.

Although the use of a control group, who would also be subjected to both a pre-test and post-test assessment, was considered in this project, the pragmatics of randomly selecting a true control group, and then putting them through pre-test and post-test assessments proved to be logistically impossible. While this unfortunate practicality weakens the research design, and therefore limits the conclusions that can be made, significant effort was made to recognise and consider jeopardising factors (146)(149)(147). The impact of history, maturation, testing, and the effect of autonomous changes in the measuring instrument or tool between the pre-intervention test and the post-intervention test may have all had an effect on the results and so these jeopardizing factors need to be considered in the conclusions. Chapter 4 describes these in the context of the pre-test quantitative assessment and also picks up the rationale behind using a comparator group of trainees from multiple year groups in an attempt to offset some of these problems.

This type of experimental design also creates issues with external validity due to what could be described as 'interaction effects'. These are interactions between the intervention and the selection process used to select the cohort group, and between the intervention and the testing

processes. This was a consequence of attempting to use the initial focus groups as a way of feeding in the trainees' views into developing the intervention itself. Although this brings in important principles of co-design (159)(160) it does have the effect of compromising interpretation of the evaluation. These factors that jeopardize external validity can lead to difficulties with generalisation of the results to other settings.

Having summarised some of the main discussion areas around the establishment of the trainee committee, detailed the implementation of the leadership development intervention and described the thinking behind the evaluation methodology, the next chapter describes the work to develop a quantitative measurement of the baseline (pre-test or 'starting point') leadership capabilities of the trainee committee cohort.

Chapter 4 – Quantitative measurement of the starting point

Chapter Overview

This chapter describes the quantitative methodologies used to measure the pre-intervention baseline (pre-test), or 'starting point'. It introduces the 'comparator' group of trainee paediatricians against which the trainee committee cohort could be compared, and describes the development of a self-assessment tool that was used for this measurement. The chapter explores the use of vignettes within the self-assessment tool, the attainment of content validity, the development of an ability scale and summarises some important issues around self-assessment. The chapter details the results of these measurements and discusses the conclusions that can be drawn from them.

4.1 Introduction

The previous chapter described the trainee committee cohort and the processes involved in refining and delivering the leadership development intervention. It also described, as an overview, the development of an overall evaluation strategy for the programme, and the rationale for the evaluation methodologies chosen. This chapter looks at one aspect of this strategy, the quantitative measurement of the pre-test baseline, or 'starting point', using a specifically designed self-assessment tool, in much greater detail.

Medical education is dominated by research that explores different methods of assessing performance or competence. A significant amount of this work describes objective assessments or examinations that relate to learning outcomes, carefully blue-printed and standard set and implemented in formal reproducible settings. Within this thesis an assessment was needed that could detect any post-test changes that might be present.

4.1.1 Finding an objective measure

Careful consideration was given to the possibility of developing an objective assessment that the cohort of trainees could sit both prior to and following the intervention. The work described in Chapter 2, which had explored the evaluation methodologies of the papers that had been included within the systematic review of the medical leadership literature, was an important source of direction. I also tried to contextualise this thinking in my learning from the leadership theories described in chapter 1. Logistically the simplest approach would have been

to have set a knowledge test that could have been implemented, perhaps even online, before and after the intervention. This was the approach taken by many of the studies found in the literature review (85)(86)(87); it was also considered in this research, but ultimately rejected. Looking at many of the competencies in the Medical Leadership Competency Framework (MLCF) (19), and reflecting on the discussions around transformational (21), distributed (28) and shared (30) leadership, the focus of this research was around development of leadership qualities, behaviours and competencies. With these intended outcomes it can be argued that a knowledge-based assessment would have relatively little meaning (56). While there is value in individuals being able to understand leadership theory, which can then go on to inform their practice, the MLCF itself uses leadership evidence-base as a framework from which to outline leadership behaviours and competencies. As discussed in Chapter 3 the leadership development intervention in this study was therefore not designed to specifically deliver knowledge, but to open up perspectives and encourage opportunities to get involved with projects and work that it was hoped would lead to learning and development. If levels of knowledge in the trainee cohort were shown to have increased in the post-intervention assessment, my opinion is that it would be difficult to interpret what that actually means.

Concepts around the assessment of competence are perhaps best presented by Miller, a psychologist, who in 1990 described four levels of competence that could be assessed in different ways (see Figure 4.1)(161). In every step up the pyramid the underlying level is a building block up to the next level. At the base of the pyramid is ‘knows’, then ‘knows how’, then ‘shows how’, and finally ‘does’:

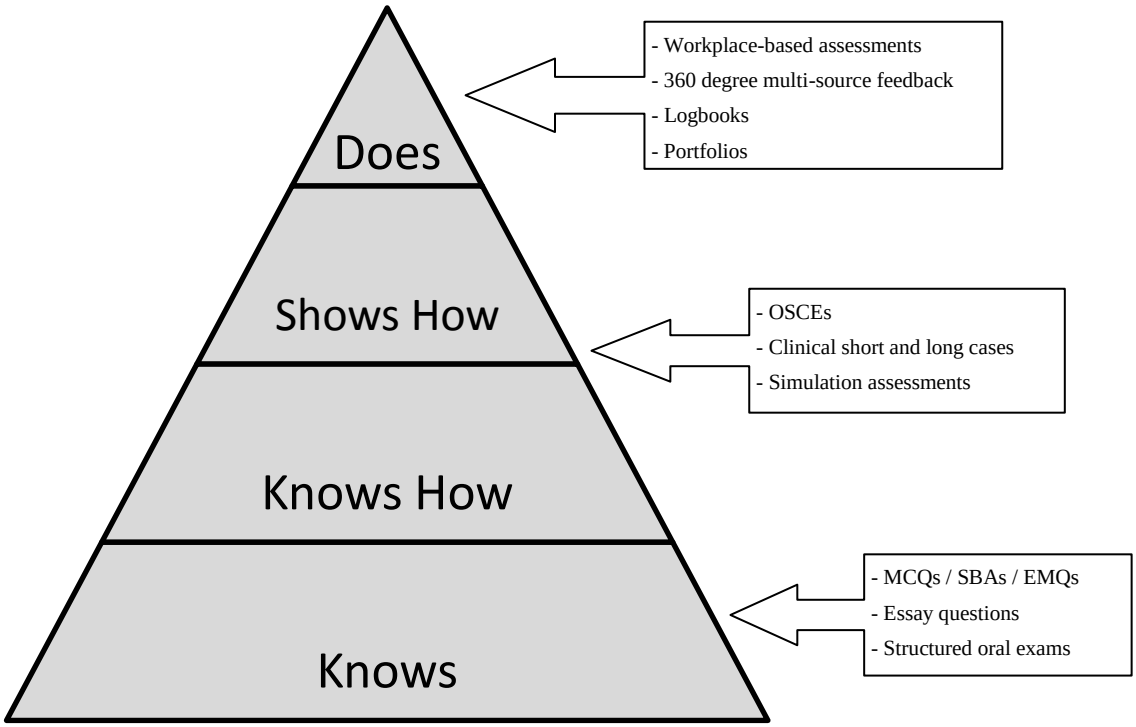


Figure 4.1 Miller's framework of clinical assessment (161)

Miller argues that examination papers, that rely on essay questions or multiple choice questions (MCQs), single-best answer questions (SBAs) or extended matching questions (EMQs) assess competence at the level of knowledge, so with the rationale described above would not be appropriate for this study. Another possibility that was considered was to develop a specific Objective Structured Clinical Examination (OSCE) (162)(163), focused on objectively scoring performance in leadership related tasks. Where OSCE questions are carefully crafted they are able to assess what the candidate *shows* they can do, but this may not correlate to what the candidate actually *does* in their everyday practice. Although there is some potential gain here, it was felt in the context of this study, that the complex logistics, in terms of design and operational implementation (expense, staff, examiners, multiple circuits, role-players), of a 'before' and 'after' OSCE out-weigh the potential benefits.

The top of Miller's pyramid is all about assessing exactly what people do, for real, in their everyday practice. In essence it is about assessing performance; how the actions and behaviours of an individual match up against expected standards. The really difficult aspect of this is coming up with metrics and tools that objectively observe and measure this. Whilst 360 degree multi-source feedback gives some summary views on performance and competence, unless one is permanently followed and watched by an objective observer, measures are reliant on self-assessment (such as portfolios, log-books and self-assessment components of multi-source feedback). This is an area, described in the sub-section below, that has considerable complexity.

4.1.2 Self-assessment

Whilst using a self-assessment strategy is very appealing in terms of logistics, there is good evidence of limitations when this process is used to assess, or score, the abilities or performance of medical professionals (164)(165)(166). The suggestion is that doctors are likely to over-estimate their self-assessment scores when compared to existing objective, externally generated markers of performance. There is also evidence to suggest that the least competent are also the least able to self-assess accurately, often significantly over-estimating their own level of performance (167).

There is, however, evidence that the accuracy of self-assessment tools can be enhanced by feedback, and also by providing explicit assessment criteria and guidance on what level of performance should score what mark (ie benchmarking) (168). There is also an argument that

self-assessment can enhance a culture of feedback and strong reflective practice (164)(168). Many successful leadership development programmes use self-assessment tools and instruments, such as the Myers-Briggs Type Indicator (MBTI) psychometric tool (121)(104), to support participants in developing further insights into their individual strengths and areas for development.

As detailed in Chapter 2, although many of the studies that had been exploring leadership development initiatives used pre- and post-intervention self-assessments, there were no useable, previously validated leadership self-assessment tools described in the literature. Of the included studies Osborn et al (83) had used a 1-5 abilities scale where 1 indicated unsatisfactory, 3 good and 5 excellent, although there appeared to have been little attempt made to validate the scale. Kouzes and Posner's Leadership Practices Inventory (LPI)(169), which was developed through mixed-methods studies across a broad range of businesses and enterprises, is an example of a validated tool. Interviews, case studies and observations of exceptional leadership were used to define the five practices of exemplary leadership. These practices formed a framework against which the LPI could be based. Broadly applicable to healthcare it has previously been used in a study on leadership in nursing (170), but because of the language used in the tool was judged to not be sufficiently applicable to this work.

Therefore, there was a need to develop an original tool for use in this project. Building on the concepts described above, where the importance of providing explicit criteria within the context of the self-assessment tool was highlighted, led to exploratory thinking about:

a) the use of vignettes to provide specific examples against which to self-assess abilities; this is explained in more detail in sub-section 4.1.3 immediately below.
b) achieving high content validity for the tool; content validity refers to the extent to which a specific measure represents all of the wider aspects of a particular social construct. The description as to how this was achieved, and the relationship back to the Medical Leadership Competency Framework, is presented in sub-section 4.3.1 below.
c) looking into triangulating self-assessment scores against a log of previous experiences. Triangulation comes from the military, or navigation, description of using multiple reference points, from different angles, to identify the exact location of an object. In this situation, it refers to using more than one methodology to explore the same phenomenon (147). The supposition in this example was that evidence of a significant number of other leadership experiences might justify a particularly high self-assessment score of ability in that particular domain. The methods used to achieve this are described in detail in sub-section 4.3.3.

4.1.3 Vignettes

There is a social sciences literature from the last 30 years describing the use of vignettes to support robust data collection within qualitative research (171)(172)(173). Vignettes have been described in many different ways, but are in essence brief scenarios on which participants can be asked to comment, and from which researchers can explore the subjective belief systems of those who have participated. Although the methodologies employed are very varied, much of the research to date using vignettes has focused on exploration of the attitudes and beliefs of participants in surveys, interviews or focus groups. The opportunity to depersonalise stories and increase both context and reproducibility of the scenario are of great value in these situations.

There are also a few examples where vignettes have been used as anchoring statements, alongside Likert-scales or other self-assessment measures, with the aim of trying to ensure answers given are more objective. This approach was demonstrated very neatly by some statistical modelling undertaken by van Soest et al who developed a model in which objective measurements were used to validate a series of vignette-based responses (174). This particular model was illustrated with reference to objective and subjective self-assessments of drinking behaviour by students in the Republic of Ireland, and the results clearly demonstrated the value of vignettes in increasing objectivity.

Within the medical literature, vignettes have been used much less, although there are increasing examples of the use of clinical vignettes to measure the competence of physicians (175) and to characterise the variations in practice between different clinicians (176). In this study the aim was to build on this work by using two brief vignettes alongside an opening descriptive statement, for each of the 11 dimensions of the self-assessment survey. An example of this, for the ninth dimension 'Planning / Applying Knowledge & Evidence', is given below:

> Doctors show leadership by actively contributing to plans to achieve service goals. They gather information to produce an evidence-based challenge to systems and processes in order to identify opportunities for service improvements. For example:
>
> (Vignette 1) imagine a situation where your consultant asks you to think about ways in which you could gather feedback from patients to help develop plans for improved adolescent services within your department.
>
> (Vignette 2) reflect on situations where you have used your knowledge of the latest evidence to challenge existing practices and processes within your department.

These pairs of vignettes are then followed by a question that asks participants:

> Reflecting on your own practice, and thinking about the two examples given above, please rate your own ability in planning and applying knowledge & evidence from 1 to 5 on the scale below:

As with the research highlighted in the literature, the aim is for the vignettes to help respondents to better understand the context of the question, to reflect on and explore previous experiences and to increase objectivity. Further description of how the vignettes were created and piloted can be found in sub-section 4.3.1 below.

4.2 Objectives

The main objective of this chapter is to describe a self-assessment tool to measure the leadership capabilities of the trainee committee cohort prior to the implementation of the leadership development intervention (pre-test), and to compare these results with a comparator group of trainees. Splitting this core project objective into smaller steps results in there being five chapter-level objectives, which are to:

- Describe the design and development of a self-assessment tool to assess trainees' leadership abilities
- Describe the methodologies used to implement this assessment with the trainee cohort and also the comparator group
- Explore whether there were any significant differences in self-assessment results between the trainee cohort group and the larger comparator group
- Explore whether the recording of previous experiences has any predictive value of self-assessed 'ability scores'
- Test whether there is any historical evidence of progression of learning as trainees moved through their clinical postgraduate training and became more senior

4.3 Methods

This section details the development and implementation of the self-assessment tool used to measure the starting point, or baseline of leadership ability, of the trainee committee cohort, prior to them commencing on the leadership programme. In addition this tool allowed for comparison of the trainee committee cohort with their paediatric trainee peers.

4.3.1 Development of the self-assessment tool

With no previously validated and published self-assessment tool available, the tool used in this study had to be developed.. This involved a number of important stages which are outlined in this sub-section.

4.3.1.1 *Content validity*

Content validity refers to the extent to which a specific measure represents all of the wider aspects of a particular social construct, in this case 'clinical leadership'. In order to establish content validity the self-assessment tool developed for this project was based on the five domains of the Academy of Medical Royal Colleges and NHS Institute of Innovation and Improvement's Medical Leadership Competency Framework (MLCF) (19) – see Figure 4.2. The MLCF was developed specifically for doctors within the NHS, and is underpinned by international evidence around leadership competency and capability in healthcare (177)(112).

Figure 4.2 Domains of the Medical Leadership Competency Framework (19)

Each of the five domains of the MLCF has four elements, making a total of 20 elements.

MLCF Domains:	**MLCF Elements:**	**Project 'Dimensions':**
Demonstrating Personal Qualities	Developing Self Awareness	(A) Developing self awareness
	Managing Yourself	(B) Managing yourself
	Continuing Personal Development	(C) Continuing personal development
	Acting with Integrity	(D) Acting with integrity & managing people
Working with Others	Developing Networks	(E) Developing networks and building & maintaining relationships
	Building & Maintaining Relationships	
	Encouraging Contribution	(F) Working within teams and encouraging contribution
	Working within Teams	
Managing Services	Planning	(G) Planning & applying knowledge/evidence
	Managing People	(D) Acting with Integrity & Managing People
	Managing Resources	(H) Managing resources and performance
	Managing Performance	

Improving Services	Ensuring Patient Safety	(I) Critically evaluating, ensuring patient safety and facilitating transformation
	Critically Evaluating	
	Facilitating Transformation	
	Encouraging Improvement & Innovation	(J) Identifying the contexts for change and encouraging improvement & innovation
Setting Direction	Identifying the Contexts for Change	
	Applying Knowledge and Evidence	(G) Planning & applying knowledge/evidence
	Making Decisions	(K) Making decisions and evaluating their impact
	Evaluating Impact	

Table 4.1 Domains and elements of the Medical Leadership Competency Framework (MLCF) (19) **and their relationship with project 'Dimensions'**

The initial version of the self-assessment tool was based on individual questions relating to each of the 20 elements of the MLCF. However, piloting with 7 paediatrician colleagues demonstrated that 20 questions were perceived by the users to be too many to be able to complete. Therefore, for the development of this self-assessment tool, these elements, which are outlined in Table 4.1, were combined into a series of 11 descriptive statements, or dimensions. This process was led by the lead researcher (RK) but was also discussed with three other senior members of the Enhancing Engagement in Medical Leadership project team, who developed the original Medical Leadership Competency Framework. These discussions reviewed the content of the 11 collapsed dimensions, and concluded that although there may be a loss of coverage and therefore content validity, the practicalities of ensuring that users felt comfortable to complete the self-assessment tool outweighed this. Each of these 11 dimensions were placed alongside pairs of examples, as described in sub-section 4.1.3 above on 'vignettes', to help contextualise them.

The vignettes were created by the lead researcher (RK) by adapting the case studies within the main MLCF (19) text as a framework for writing paediatric-specific vignettes that would resonate with the trainee committee cohort and paediatric trainee comparator groups. Relating this directly back to the MLCF was an important way of ensuring that the full breadth of each of the competencies (eg self-awareness) was incorporated. Feedback on the readability and applicability of the vignettes was obtained during the piloting phase and adaptations were made before a final version was confirmed.

Within the self-assessment tool these combinations of statements and examples were used as a stimulus to ask participants about their self-perceived ability in each area. All of the statements and examples can be found within the final version of the self-assessment tool in Appendix 4.1.

The MLCF and the wider NHS 'Leadership Framework' which followed from it in 2011(158) are increasingly widely used across the NHS and internationally. Interestingly, since completing my work in this area, a Leadership Framework Self Assessment tool has been published by the NHS Institute of Innovation and Improvement. The aim of this tool, which although similar in content has a different scale and approach to the tool developed for this project, is to support NHS workers to manage their own leadership learning and development (157).

4.3.1.2 *Developing an 'ability scale'*

Having designed and piloted the 11 descriptive statements and pairs of examples for the self-assessment tool, a measurement scale, consisting of a number of clear criteria against which an individual could assess themselves, was developed. This process took into account a number of key considerations including whether or not to use pure Likert scales, or Likert-*type* scales such as semantic differential scales (178)(179), the number of options, the words and/or numbers used in the scale and the nature of the data produced.

There is significant debate within the medical education and social sciences literature as to whether Likert scales, 'Likert'-type scales, or semantic differential scales (178), should have an odd or even number of points, and indeed how many points there should be (180)(181). The concern about an odd number of points is that a disproportionate number of respondents head for the neutrality of the middle option. However, with an even number there is no possibility of a 'middle option', which is perhaps unrealistic, and for that reason a 5-point scale was adopted in this study.

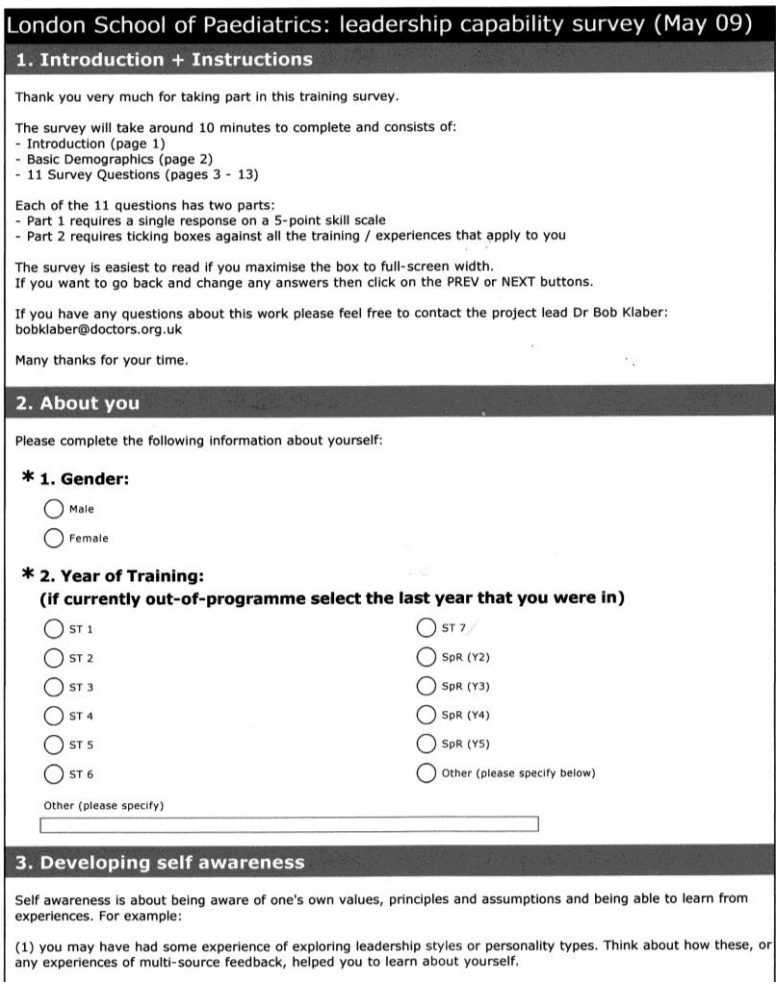
Another much argued issue is whether Likert scales, and Likert-type scales produce interval or ordinal data (182)(180). This, along with other factors around the assumed normal distribution of the data, has important implications as to whether parametric tests, such as a t-test (likely to be more statistically powerful than its non-parametric equivalent) can be used in analysing the data. To strengthen the case for the data being interval, a number scale from 1 to 5 was evenly transcribed between the two descriptors used in the scale. The assumptions around this data and the choice of statistical methods used are discussed in sub-section 4.3.4.

Options were explored about adapting the widely used Strongly Disagree – Disagree – Neither Agree nor Disagree – Agree – Strongly Agree scale often seen in traditional 5-point Likert scales, or in adapting a previously published 'confidence scale'. Neither of these would have

achieved the desired assessment of ‘readiness’ or ‘preparedness’ so after trialling a number of drafts, it was decided to use a semantic differential scale, with bipolar descriptors at either end of a five-point scale. This adapted five-point ‘skills scale’ was used as follows:

(1) I find this difficult and I have to consciously work at it	(2)	-	(3)	-	(4)	-	(5) This is so inherent in my practice that I find it easy and instinctive

This scale, based on the principles of Dreyfus model of skills acquisition (183), was adapted from one used in a small qualitative study exploring confidence and competence (184). The two statements describing skills levels were aligned to the numbers 1 and 5, with the numbers 2,3 and 4 spaced evenly between them, in order to bolster the argument of this being an interval scale. This scale was pre-piloted with a group of 8 trainees from a different specialty to check that it could be quickly read and clearly understood. This gave a further opportunity to pilot the format and readability of the vignettes. One improvement to the layout of the scale was made following this pre-pilot. The content and formatting of the final scale used can be seen within the first three sections of the Portable Document Format (PDF) of the final version of the self-assessment tool in Figure 4.3 below, or in full at the end of the thesis in Appendix 4.1.



London School of Paediatrics: leadership capability survey (May 09)

1. Introduction + Instructions

Thank you very much for taking part in this training survey.

The survey will take around 10 minutes to complete and consists of:
- Introduction (page 1)
- Basic Demographics (page 2)
- 11 Survey Questions (pages 3 - 13)

Each of the 11 questions has two parts:
- Part 1 requires a single response on a 5-point skill scale
- Part 2 requires ticking boxes against all the training / experiences that apply to you

The survey is easiest to read if you maximise the box to full-screen width.
If you want to go back and change any answers then click on the PREV or NEXT buttons.

If you have any questions about this work please feel free to contact the project lead Dr Bob Klaber: bobklaber@doctors.org.uk

Many thanks for your time.

2. About you

Please complete the following information about yourself:

*** 1. Gender:**

- ◯ Male
- ◯ Female

*** 2. Year of Training:**
(if currently out-of-programme select the last year that you were in)

- ◯ ST 1
- ◯ ST 2
- ◯ ST 3
- ◯ ST 4
- ◯ ST 5
- ◯ ST 6
- ◯ ST 7
- ◯ SpR (Y2)
- ◯ SpR (Y3)
- ◯ SpR (Y4)
- ◯ SpR (Y5)
- ◯ Other (please specify below)

Other (please specify)

3. Developing self awareness

Self awareness is about being aware of one's own values, principles and assumptions and being able to learn from experiences. For example:

(1) you may have had some experience of exploring leadership styles or personality types. Think about how these, or any experiences of multi-source feedback, helped you to learn about yourself.

(2) imagine a situation where your own emotions and feelings could affect your judgement and behaviour. Consider how you might manage the effect of stress on your behaviour.

*** 1. Reflecting on your own practice, and thinking about the two examples given above, please rate your own ability in developing self awareness from 1 to 5 on the scale below:**

	(1) I find this difficult and I have to consciously work at it	- (2) -	- (3) -	- (4) -	(5) This is so inherent in my practice that I find it easy and instinctive
Tick one:	◯	◯	◯	◯	◯

*** 2. What experiences / training have helped you to develop self-awareness?**

Please tick all that apply:

- ☐ No experiences / training
- ☐ Experiences at school
- ☐ Experiences at university
- ☐ Learning from role-models
- ☐ Experiences with patients
- ☐ Educational supervision
- ☐ Multi-source feedback
- ☐ Trust-based teaching sessions
- ☐ Reading books / journals
- ☐ Scenario-based teaching (eg APLS)
- ☐ Courses / conferences
- ☐ Multi-disciplinary meetings
- ☐ Clinical audit projects
- ☐ Attending management meetings
- ☐ 'Improvement' projects
- ☐ Contributing to national working groups
- ☐ Contributing to patient safety initiatives
- ☐ Teaching / supervising colleagues
- ☐ Experiences from undertaking research
- ☐ Other experiences (see text box below)

Other experiences (please specify):

Figure 4.3 First 3 sections of pre-intervention self-assessment questionnaire

4.3.1.3 *Evidencing self-assessment scores*

As described in sub-section 4.1.2, there is good evidence of the limitations of using a self-assessment approach in isolation (164)(166), so a strategy was developed to correlate self-assessment scores against a log of previous experiences. The aim was to strengthen the design by using a second methodology to explore the same phenomenon. The questionnaire was designed such that after making their self-assessment of leadership ability, for each of the 11 leadership areas, participants were also asked to illustrate which previous experiences they had used to help gain relevant knowledge, skills, attitudes and behaviours. They did this by checking the boxes of each of a long list of possibilities (which included an 'Other' option with a free text box) to record which ones they had previously experienced.

The aim of this process was to add to the validity of the self-assessment by seeing whether there were areas of knowledge, or groups of participants, where the level of self-perceived ability was not underpinned by any significant and relevant experiences (168). This would potentially cast doubt on the validity of the self-assessment scores as logic would suggest that a position where one "finds something easy and instinctive" would be underpinned by significant prior experience. In addition, it was decided to use a process of multivariate analysis of variance, to explore whether any of the experiences that people may have had were predictive of significantly higher scores compared to people who had *not* had those experiences, in one or more of the eleven dimensions of leadership assessed. The results of this analysis are discussed in sub-section 4.4.5, and critique of some of the inherent assumptions around this are detailed in section 4.6.

4.3.2 Piloting

The self-assessment questionnaire was piloted, initially in paper form, and then subsequently using the web-based format in which it was finally distributed. Feedback from 7 doctors (mixture of trainees and consultants) was received, and an additional 8 doctors from a different specialty commented on the five-point scale. This feedback informed adaptations to the format and language used in order to improve the clarity and flow of the questionnaire. This included feedback on the terminology used, which helped to simplify and clarify the descriptions of the context and the assessment instructions. One significant concern, despite the decision to reduce the number of dimensions down from 20 to 11, was around the length of the questionnaire. There are a number of studies that demonstrate a falling response rate, and increased numbers of incomplete questionnaires, the longer the survey is, although much of the work on this has been done on postal questionnaires (185)(186)(187). There are also potential concerns about

answer fatigue reducing the quality of responses, although this is not borne out in the literature (185).

As highlighted above, it was for these reasons that the original design of the questionnaire was reduced from 20 questions (corresponding to the 20 elements within the MLCF) to 11 questions, with elements being condensed and combined together to form 11 dimensions (see Table 4.1). It was recognised that this has some potential impact on the content validity and discussion around this is picked up in section 4.6. Even with these improvements, piloting demonstrated that the average time of completion was just over 18 minutes, which it was acknowledged was on the longer side of completion time for the breadth of surveys commonly received by doctors. The risk here is that respondents may become fatigued or bored and either rush their answers, or leave the survey without completing it.

4.3.3 Implementation

The self-assessment questionnaire was designed and formatted on a widely used, internationally available, web-based survey platform (www.surveymonkey.com). A postal survey was considered, as there is some evidence that response rates through this method are higher than via web-based surveys (eg 21% v 11% in a survey of 6000 US dentists (188)), but for economic, implementation and data analysis reasons the web-based approach was chosen.

The self-assessment questionnaire, once designed and piloted, was sent out to the trainee committee cohort, and the full group of comparators (every other paediatric trainee within the London School of Paediatrics and Child Health) via individual e-mails. This technique, where the main introductory message and survey URL link was pasted into the body of each e-mail to each individual potential participant, allowed for personalisation of the request to complete the survey. Although difficult to know whether this had any impact, within the context of a group of recipients who receive anywhere between 5 and 20 web-based surveys per annum, and despite being labour intensive without the support of mail-merge technology, this seemed to be a worthwhile approach to personalise the request with the aim of improving the response rate.

A single, individual e-mail reminder was sent to potential participants between 21 and 28 days after the original request. The survey remained open for a further 6 weeks although the majority of the responses came within 48 hours of sending the requests.

4.3.4 Statistical analysis and rationale

Having designed a methodology where a questionnaire asked participants to illustrate previous experiences that had supported the development of their self-assessed leadership abilities, a number of different approaches were explored as means of analysing potential correlations, and in determining any significant underlying factors.

The data from the web-based self-assessment questionnaires was initially downloaded into Microsoft Excel format, before being manually cleaned in preparation for being converted to SPSS format. The analysis was performed within SPSS; support and guidance for this and input into the choice and running of statistical tests came from an experienced medical statistician, Henry Potts, who is an academic colleague within University College London.

The approach to the statistical analysis of the self-assessment data aimed to begin by exploring whether there was evidence of global changes or differences between data (eg using an ANOVA test). There is then an argument that only if these global differences are found does the analysis move to a more specific level of detail, for what are often called 'post-hoc' tests (eg in this case using T-tests to explore each of the different dimensions). As this project has a mixed methods approach, relating this analysis back to the original questions and the data gathered from the qualitative work also helped to give the findings important context and interpretation.

As described above in sub-section 4.3.1.2 there is some contention as to whether the data from Likert-type scales, or semantic differential scales, is interval or not. This has important implications as to whether or not parametric tests can be used. For this project histograms of the data were generated and reviewed within SPSS. In all cases they looked to be normally distributed, with no evidence of a long tail (this would be unlikely with only 5 data points) or skewing of the data. As a result it was concluded that it was possible to use parametric statistical tests on this self-assessment score data as they were judged to be sufficiently robust. Where the data was much more likely to be skewed (eg with the experience log data) non-parametric tests were used. The sub-sections below describe the different tests used.

4.3.4.1 *Analysis of variance*

This 'global' test explores whether there is any change or difference between the two sets (ie cohort and comparator, or in the case of Chapter 6 pre- and post-) of data. Even if none of the 11 questions were to show any *significant* difference as individual questions, there remains the

hypothesis that there might have been a 'small amount of change across all 11 dimensions', and collected together this might amount to a significant difference, so this was the rationale for running this test. This concept was analysed on SPSS using a global analysis of variance (ANOVA) test. It is worth noting that this is a parametric test, and there is no non-parametric equivalent, so the assumptions and conclusions about interval data that are made in the paragraph above are important.

4.3.4.2 *T-tests*

If there was a suggestion from the global analysis that there might be some changes or differences to try to explain, the self-assessment scores of the different groups, and the scores for each of the 11 different dimensions were compared using unpaired 'independent samples' t-tests to look for any significant differences. The pre- and post-intervention scores of the trainee committee cohort were analysed using a paired t-test (see Chapter 6). T-tests are a parametric test, but there is a non-parametric equivalent test that could be used if the assumptions about the interval nature of the semantic differential scales, discussed in sub-section 4.3.1.2 above, were ruled to be invalid.

4.3.4.3 *Bonferroni correction*

Bonferroni corrections (189) can be applied to account for the relatively small samples sizes. This adjustment entails reducing the critical level of 'significance' from 0.05 by dividing it by the number of tests being run. With the 11 different questions within the self-assessment questionnaire, and the data arising from each, this would lead to a revised cut off for 'significant' of 0.05 divided by 11, which is 0.0045. This correction, which is widely viewed as being quite conservative, means that there have to be fairly large differences in scores between groups, or between the pre- and post-intervention results, to attain significance. There is an argument that in the case of post-hoc tests, being run after obtaining positive global test results, results can legitimately be displayed without applying the Bonferroni correction. The implications of whether or not the correction is applied are described in the results section.

4.3.4.4 *Factor analysis & dimension reduction*

A process of factor analysis, using SPSS, was applied to explore whether or not the 11 different questions in the self-assessment could be statistically grouped together into a significantly fewer number of larger dimensions.

The rationale behind this choice was to take a more analytical approach to collapsing down the full breadth of the MLCF elements, to see if a small number of interpretable factors emerge. The hope was that this might further inform understanding about the areas of learning need for trainees of different levels, and hence help to shape leadership development interventions in this area. This process of factor analysis is also known as dimension reduction and is a well-recognised method of combining multiple strands of observed data into a fewer number of non-observed dimensions which can then be used to describe results (190).

4.3.4.5 *Pearson's product-moment correlation coefficient*

Pearson's product-moment correlation coefficient is a parametric test that can be used to compare the statistical dependence of two variables. It has a non-parametric equivalent test called Spearman's rank correlation coefficient, but because of the arguments around the interval nature of the semantic differential, or Likert-type, data, a decision was made to use Pearson's test. In this study it was used on the pre-intervention data (for both the trainee committee cohort and the wider comparator group) to see if there was any association between the numbers of years of training an individual had received and the self-assessment scores in each of the 11 domains. The rationale for this test was to explore, in the absence of a post-test assessment of the comparator groups, whether or not there was evidence that trainees naturally developed leadership skills and behaviours as they went through the years of the standard postgraduate paediatric training programme. The value of this additional experiment, and what can be concluded from it, is discussed in section 4.6. An assessment as to whether these results were any different using 'grade of training' rather than the number of years in training was also made.

4.4 Results

This section describes the quantitative results of the work to measure the starting point (pre-test), generated by the self-assessment tool described above.

4.4.1 Response rates

The URL link to the self-assessment tool was sent by e-mail to 919 paediatric trainees as described in sub-section 4.3.3. Responses were received by 336 trainees (37% response rate); of these 336 trainees, 289 fully completed the assessment (86% completion rate). This gives an amended response rate for fully completed self-assessment questionnaires of 31%. In addition to these numbers, of the 32 doctors in the trainee committee cohort, 27 (84% response rate) fully completed the pre-intervention self-assessment questionnaire. Table 4.2 shows the respondents from each group according to their training grade:

Training Grade	Comparator Group (pre-test)	Trainee Committee Cohort (pre-test)
ST1	35	3
ST2	41	6
ST3	32	2
ST4	54	4
ST5	42	3
ST6	36	5
SpR (ST7)	47	4
SpR (ST8)	47	1
Post CCT	2	0
Total	336	28

Table 4.2 Number of self-assessment questionnaire respondents by training grade

The comparator group had between 1 and 8 years of specialty training in paediatrics, with a median of 5 years (IQR 3-7). The trainee committee cohort also had between 1 and 8 years of specialty training in paediatrics, with a median of 4 years (IQR 2-6). The self-assessment scores for each individual within the trainee committee cohort and the comparator group are

illustrated in Appendix 4.2. Due to the size of spreadsheet, the raw 'experiences' data has not been included.

4.4.2 Comparison of trainee cohort and comparator group

For the comparator group to add any value to the experimental methodology it was necessary to determine that there were no significant differences in terms of demographics and self-assessed leadership ability between them and the paediatricians in the cohort group. With the trainee committee cohort being a self-nominated group there was the possibility that their self-assessed leadership ability may have been above average.

In order to test for these differences, a repeated measures ANOVA was performed with a within-subjects factor of the 11 dimensions and a between-subjects factors comparing the cohort group and the comparator group. This 'global test' showed no significant difference when comparing the self-assessment scores of the two groups ($F_{1,314} = 0.4$, $p = 0.5$) and no significant interaction between group and dimension ($F_{10,305} = 0.6$, $p = 0.8$).

Although the 'global test' described above was able to reassure that there was no statistical difference between the self-assessment scores of the cohort and comparator groups, this concept was explored in greater detail by looking at each of the 11 dimensions individually. The possibility of there being significant differences between the mean scores of the cohort and comparator groups for each of the 11 dimensions were calculated for both groups and compared using unpaired t-tests. The rationale for this choice of statistical analysis is described in sub-section 4.3.4.1.

The results, which have been collated in Table 4.3, clearly show no significant difference between the self-assessment questionnaire scores of the two groups across all 11 dimensions. There is therefore no evidence that, prior to the introduction of the intervention, the study group differed significantly from the comparator group on any of the 11 dimensions. This correlates with the findings of the repeated measures ANOVA global test described above.

Dimensions	Comparator group: mean	N	Trainee committee cohort: mean (N = 27)	Comparison of comparator group with trainee committee cohort
(A) Developing self awareness	3.4	336	3.5	$t_{361} = -0.3,\ p = 0.8$
(B) Managing yourself	3.6	328	3.7	$t_{353} = -1.1,\ p = 0.3$
(C) Continuing personal development	3.6	317	3.7	$t_{342} = -0.7,\ p = 0.5$
(D) Acting with integrity and managing people	3.4	314	3.4	$t_{339} = 0.1,\ p = 0.9$
(E) Developing networks and building & maintaining relationships	3.3	310	3.3	$t_{335} = -0.5,\ p = 0.6$
(F) Working within teams and encouraging contribution	3.7	306	3.6	$t_{331} = 0.8,\ p = 0.4$
(G) Planning and applying knowledge & evidence	3.1	303	3.3	$t_{328} = -1.0,\ p = 0.3$
(H) Managing resources & performance	3.3	302	3.2	$t_{327} = 0.2,\ p = 0.8$
(I) Critically evaluating, ensuring patient safety and facilitating transformation	3.2	299	3.1	$t_{324} = 0.5,\ p = 0.6$
(J) Identifying the contexts for change and encouraging improvement & innovation	3.2	294	3.3	$t_{319} = -0.9,\ p = 0.4$
(K) Making decisions and evaluating their impact	3.0	289	3.2	$t_{314} = -1.1,\ p = 0.3$

Table 4.3 Comparison of pre-intervention self-assessment scores between the comparator and trainee committee groups.

Table 4.3 also illustrates the drop-off in numbers of answers throughout the 11 dimensions of the self-assessment questionnaire. After an initial drop-off where 8 respondents only answered the first question, and a further 11 only got as far as the second question, the drop-off remained fairly even throughout the questionnaire. The calculations were made independently for each dimension using the total number of respondents to that specific dimension. This meant that for this analysis the partial respondents were fully included. All 27 of the trainee committee cohort who started the questionnaire completed it.

Finally, in order to examine potential differences in answers from female and male participants, another repeated measures ANOVA, this time with a between-subjects factor of gender, was performed. This showed that there is no main effect of gender ($F_{1,314} = 1.4$, $p = 0.2$) and no interaction with dimension ($F_{10,305} = 1.4$, $p = 0.2$), meaning that the effect of gender was not particularly significant in this work.

4.4.3 Using factor analysis to determine summary themes

As described in sub-section 4.3.4.4 an exploratory factor analysis was performed to explore whether there were any factors that could encompass more than one of the 11 dimensions set out in the original self-assessment tool. This was done using SPSS.

The factor analysis demonstrated a reduction from the 11 observed leadership dimensions down to 2 significant factors. The first factor was able to explain 27% of the variance, and the second factor, an additional 20%, reaching almost 50% between these first two factors. In the analysis, weights over 0.3 were considered to be significant, and all of these are illustrated in Table 4.4. Where weights were under 0.3 they have not been displayed in order to avoid distracting 'statistical noise'.

Dimensions	Factor 1 (weight)	Factor 2 (weight)
(A) Developing self awareness		0.73
(B) Managing yourself		0.58
(C) Continuing your personal development		0.60
(D) Acting with integrity and managing people		0.58
(E) Developing networks and building & maintaining relationships	0.67	
(F) Working within teams and encouraging contribution	0.46	0.36
(G) Planning and applying knowledge & evidence	0.65	
(H) Managing resources & performance	0.52	0.43
(I) Critically evaluating, ensuring patient safety and facilitating transformation	0.63	
(J) Identifying the contexts for change and encouraging improvement & innovation	0.81	
(K) Making decisions and evaluating their impact	0.67	

Table 4.4 Results of dimension reduction using factor analysis

4.4.4 Progression through training

The relationship between the self-assessment scores within each of the 11 dimensions and the number of years each of the respondents had spent in specialty training was also explored. In case (usually due to time out of programme or flexible working) it was any different from 'number of years in training', the relationship was also looked at per 'grade of training'. These analyses used the data from all 363 (336 comparator group and 27 trainee committee cohort) respondents. Determining this relationship was important as it would help to explore if there were increases in scores with seniority. If this was the case, this would suggest that trainees, without the benefit of the leadership development initiative, increase their scores over time. Conversely, if there was no significant increase in scores as trainees became more senior, this could suggest that there would have been no 'natural' increase in scores over time in training within the comparator group. As described previously this was potentially important as a way of partially offsetting the lack of a post-test assessment in the comparator group in the evaluation methodology.

Pearson's product-moment correlation coefficient was used to explore whether there was any significant correlation. The rationale behind this choice is outlined in the methods sub-section in 4.3.4.5. The correlation between 'years of training' and 'grade of training' was shown to be $r = 0.99$ ($p < 0.001$) so there was no need to do the calculations for both; the actual numbers in this sub-section refer to 'years of training'. The initial results showed that there was a significant correlation across dimensions B, C and K, and this is illustrated in Table 4.5. However, if the Bonferroni correction (as described in sub-section 4.3.4.2) was applied it was only dimension (B) 'Managing yourself' ($p=0.011$) where significant correlation remained.

Dimension	Pearson's correlation (r)
(A) Developing self awareness	$r = 0.037$, $p = 0.469$
(B) Managing yourself	$r = 0.169$, $p = 0.001$
(C) Continuing your personal development	$r = 0.124$, $p = 0.017$
(D) Acting with integrity and managing people	$r = -0.004$, $p = 0.941$
(E) Developing networks and building & maintaining relationships	$r = 0.091$, $p = 0.085$
(F) Working within teams and encouraging contribution	$r = 0.006$, $p = 0.914$
(G) Planning and applying knowledge & evidence	$r = 0.061$, $p = 0.254$
(H) Managing resources & performance	$r = 0.067$, $p = 0.211$
(I) Critically evaluating, ensuring patient safety and facilitating transformation	$r = 0.083$, $p = 0.121$
(J) Identifying the contexts for change and encouraging improvement & innovation	$r = -0.009$, $p = 0.869$

(K) Making decisions and evaluating their impact	$r = 0.153$, $p = 0.005$

Table 4.5 Pearson's product-moment correlation coefficient analysis of correlation between self-assessment questionnaire scores and number of years in specialty training

With the significant result for dimension (B) the question arises as to whether there are significant periods within the duration of the paediatric training programme where the increase in scores occurs, or whether the increases are relatively evenly spaced throughout training. A regression plot was attempted with the data for this dimension, but there were insufficient numbers of data to be able make any meaningful conclusions from the plot.

4.4.5 Does the experience log correlate with self-assessment scores?

Sub-section 4.3.1.3 describes the process of triangulation, where for a specific dimension the average self-assessment scores were correlated with the average number of 'experiences' boxes that were ticked for that particular dimension. It would be expected that a position where one "finds something easy and instinctive" would be backed up by significant prior experience.

This analysis was performed for dimension (A) "Developing self awareness". Using a sample size of 263 people, who fully completed this section, the numbers of experiences counted were compared with the score for that dimension. On this occasion, because of the skewed distribution of the bimodal experiences data Spearman's rank correlation coefficient, a non-parametric test, was used to explore whether there was any significant correlation. The results showed a statistically significant correlation where $r_s = 0.26$ ($p < 0.001$). This finding offers some support to the suggestion that the 5-point self-assessment scale used in this study is a valid measure.

Having demonstrated this significant correlation the question emerges as to whether any of the previous experiences that respondents may have had were predictive of significantly higher scores compared to people who had *not* had those experiences. This phenomenon was explored using a multivariate ANOVA across all eleven dimensions of leadership assessed.

Table 4.6 details the summary results of this ANOVA through documenting the *F*-statistic (a value that arises from the statistical test that explores differences between means within an ANOVA) and p-values for each dimension. It also illustrates the 'Adjusted R squared' percentage, which gives a conservative estimate of how much of the variance in scores across

that dimension can be explained by whether or not certain previous experiences had been self-recorded as having occurred. In all 11 dimensions the ANOVA suggested that there was a significant relationship between the semantic differential scale self-assessment scores and documented previous experiences.

Dimension	*F*-statistic	p-value	Adjusted R squared
(A) Developing self awareness	$F(19,369) = 3.110$	0.000	9.4%
(B) Managing yourself	$F(19,360) = 2.013$	0.008	4.8%
(C) Continuing your personal development	$F(19,348) = 2.362$	0.001	6.6%
(D) Acting with integrity and managing people	$F(19,345) = 3.376$	0.000	11.0%
(E) Developing networks and building & maintaining relationships	$F(19,341) = 5.536$	0.000	19.3%
(F) Working within teams and encouraging contribution	$F(19,337) = 2.565$	0.000	7.7%
(G) Planning and applying knowledge & evidence	$F(19,334) = 4.201$	0.000	14.7%
(H) Managing resources & performance	$F(19,333) = 3.496$	0.000	11.9%
(I) Critically evaluating, ensuring patient safety and facilitating transformation	$F(19,330) = 4.084$	0.000	14.4%
(J) Identifying the contexts for change and encouraging improvement & innovation	$F(19,325) = 3.616$	0.000	12.6%
(K) Making decisions and evaluating their impact	$F(19,320) = 4.655$	0.000	17.0%

Table 4.6 Multivariate ANOVA detailing the relationship between semantic differential self-assessment scores and documented previous experiences

In a further stage of analysis, the potential influence of each specific individual experience on the self-assessment scores was ascertained.

Table 4.7 illustrates which particular individual experiences had a significant ($p < 0.05$) influence on increasing self-assessment scores for each of the 11 dimensions. Of note, experiences at university, learning from role-models, experiences with patients, 'improvement' projects, teaching/supervising colleagues and 'other' experiences all had a significant influence in at least 3 different dimensions. The only experience that was found to have had *no* significant influence on any of the dimensions was participating in a clinical audit project.

Dimension → **Experience ↓**	(A) Self awareness	(B) Managing yourself	(C) Continuing personal development	(D) Acting with integrity / manage people	(E) Developing networks / relationships	(F) Working within teams	(G) Planning & applying evidence	(H) Managing resources & performance	(I) Critically evaluating / transformation	(J) Context / innovation & improvement	(K) Making decisions / evaluating impact
Experiences at school				■				■			
Experiences at university							■		■	■	■
Learning from role-models	■		■	■							
Experiences with patients					■		■		■	■	■
Educational supervision	■							■			
Multi-source feedback					■						
Trust-based teaching sessions							■				
Reading books / journals			■								
Scenario-based teaching (eg APLS)										■	
Courses / conferences						■					
Multi-disciplinary meetings										■	
Clinical audit projects											
Attending management meetings					■						■
'Improvement' projects		■			■			■			
Contributing to national working groups				■		■					
Contributing to patient safety initiatives					■				■		
Teaching / supervising colleagues		■		■			■				
Experiences from undertaking research							■				
Other experiences	■			■	■					■	

Table 4.7 Experiences (marked in grey) that had a significant ($p < 0.05$) influence on increasing self-assessment scores for each of the 11 dimensions

4.5 Conclusions

This section summarises the conclusions that can be drawn from the results of the starting point survey described above. In summary these are that:

a. prior to the introduction of the intervention, the study group were not significantly different from the comparator group in terms of their self-assessed leadership abilities

b. through a process of factor analysis, the 11 different leadership dimensions could be reduced down to two significant factors. The first factor could explain 27% of the variance, and the second factor, an additional 20%
c. there is little evidence to suggest that the majority of trainees, who have not had the benefit of any specific leadership development experiences, increase their leadership abilities over time. The more senior trainees only self-assessed themselves significantly higher than more junior colleagues in one leadership dimension (B) 'Managing yourself'. This is an important conclusion in the context of there being no post-test assessment of the comparator group.
d. there is a significant correlation between the occurrence of previous experiences and the self-assessment score, which supports the validity of the self-assessment scale measure
e. certain specific previous experiences are more likely to influence the development of self-assessed leadership abilities than others

4.6 Discussion

The essence of this chapter has been to describe the work that has been done to meaningfully assess the leadership abilities of the trainee committee cohort, prior to them starting on the programme (pre-test).

As discussed in sub-section 4.3.1 a number of different approaches were taken to ensure the validity of the self-assessment measure used in this study. Basing the content and language on the domains and elements of the Medical Leadership Competency Framework, around which the lead researcher (RK) had experienced a significant amount of development and implementation work on behalf of the NHS Institute of Innovation and Improvement, gave strong content validity. There is an argument that in reducing the 20 elements down to 11 dimensions some of this strong content validity may have been diluted. As noted in the methods section, this had to be balanced against the pragmatics of using a survey that was so long that it was at risk of not being completed. With a number of mitigating steps put in place, it was hoped that the effect of this reduction would not significantly affect the results.

The use of vignettes helped to strengthen the self-assessment tool further. Their origin from within the Medical Leadership Competency Framework materials, followed by a process of piloting with a small group of paediatricians, gave them a sense of content validity as well as practical applicability. Had more time been available it may have been advantageous to have subjected these vignettes to more discussion and piloting before they were finalised.

Although it was not possible to use a previously validated scale, the detailed approach to the design of the scale, focusing on getting the structure and presentation as strong as possible, and the reference to the Dreyfus work on expertise (183) led to a measure that was likely to work reasonably well. It was also concluded that it would enable the desired assessment of 'readiness' or 'preparedness' better than a traditional Likert scale (181). There are however, on reflection, some inherent weaknesses in the scale design that might have been avoided had a Likert scale been adapted, or a different approach to creating a semantic differential scale taken (179). It can be argued that the two ends of the scale are not technically bipolar opposites, and, perhaps more critically, that the two extremes contain more than one concept (how I find it, and how I have to approach it). The scale could also be criticised for not having a clearly defined neutral mid-point:

(1) I find this difficult and I have to consciously work at it	(2)	-	(3)	-	(4)	-	(5) This is so inherent in my practice that I find it easy and instinctive

There is therefore a risk that different participants may interpret the scale in different ways, and that the data is perhaps not truly parametric in nature. This might have an impact on some of the assumptions made and discussed earlier in the methods section of this chapter, and on the results discussed in chapter 6.

The piloting process allowed for an early assessment of the functionality of the scale and the work to show that the experience log correlates with the self-assessment scoring along the interval scale offered further evidence for construct validity. The efforts, with individualised communication and follow-up e-mails, to achieve a high response rate to the surveys were also an important factor in achieving meaningful results. The impact of non-respondents, and the potential that they might have answered very differently to the respondents, is always an important consideration in survey based studies. Achieving response rates of 37% in the comparator group (with 86% of these fully completing the survey) and 84% in the trainee committee group was an important part of mitigating against this. However, almost two-thirds of the trainees in London did not respond which leads one to conclude that there is a possibility that the 'comparator' group was to an extent itself 'self-selected' in that they were all people in the one-third of trainees who felt motivated enough to participate in the survey. This discussion around non-respondents highlights another important point that was mentioned in chapter 3, namely that this self-assessment tool, which was used as one part of the pre-test assessment, is also in effect part of the intervention. This important concept, and potential jeopardising factor, could have a significant impact on the results (146)(147). Without there being a post-test comparison between the cohort and comparator groups at the end of the study

period it is difficult to properly mitigate against interference from this testing process. The pre-test focus groups described in chapter 5 generate the same issue. It may be that if this leadership intervention were to be reproduced that the self-assessment tool could also be used by individual participants to help them to reflect on their strengths and areas needing focused development (168).

Although helpful to be able to conclude that at the 'starting point', or pre-test stage, the study group were not significantly different from the comparator group in terms of their self-assessed leadership abilities, it is important to consider the possibility that no significant difference was found due to the weakness of the measurement tool. In other words, there is a possibility that there were differences between the comparator and cohort groups, but that the self-assessment measurement scale used in this study was not sufficiently sophisticated to detect them. The finding that the more senior trainees only self-assessed themselves significantly higher than more junior colleagues in one leadership dimension perhaps also backs up this suggestion, although there are two potential challenges to this assertion. Firstly, this result comes after applying a Bonferroni correction, which as described in sub-section 4.3.4.3 is widely regarded as being very conservative, and so may have 'over-corrected' any observed differences in other dimensions (two other dimensions did show some significant increase across the years prior to applying Bonferroni correction). Secondly, it is worth challenging the assumption that more senior trainees necessarily have significantly more leadership experience than their junior colleagues. Prior to 2009, when this part of the study was performed, there were relatively few leadership learning opportunities available to trainees and so most reached the end of their final year as a trainee without any specific experiences.

This data indicates that in 2009 there was little evidence to suggest that the majority of trainees, who had not had the benefit of any specific leadership development experiences, increased their leadership abilities over time. In the absence of any post-test assessment of the comparator group, as described and accounted for in chapter 3, this conclusion is an important point, and helps mitigate, to some extent, against the important jeopardising factors of history and maturation (146). Over the last 3-4 years there has been a large amount of work undertaken to embed leadership learning in postgraduate medical curricula and to increase opportunities for work-based leadership learning. So, if this part of the study was to be repeated in the future, more significant differences between the different grades of trainee might be expected to be seen.

The factor analysis that was undertaken on the self-assessment questions demonstrated the possibility of a dimension reduction down to two significant factors. These two factors, one focused on personal leadership development ('Factor 2') and the other on systems learning ('Factor 1'), align extremely well with what is known to be important from the leadership literature reviews described in chapters 1 and 2, and with the qualitative work undertaken around the time of application with the trainee committee (see chapter 3). This provides further content validity to the use of the 11 dimensions, but also opens up the idea that future work might be able to use shorter, more effective self-assessment tools, potentially with fewer dimensions that duplicate less of the same content.

Another important aspect of the self-assessment pre-test survey was the work to use documentation of previous experiences as a way of validating the answers given on the ability scale. This was based on the inference that would suggest that a position where one "finds something easy and instinctive" would be underpinned by significant prior experience. This approach was also supported by positive feedback from the individuals who participated in piloting. As described in sub-section 4.4.5 this correlation was demonstrated in a number of ways. This was initially done through using the first dimension (A) "Developing self awareness" to demonstrate a significant correlation between the number of experiences ticked and the self-assessed leadership ability score. While a helpful start, it is important to acknowledge that this result does not account for any variability in the depth, frequency, timings or quality of the experiences. Looking at the correlation from the other direction an ANOVA was used to explore whether any of the previous experiences that respondents may have had were predictive of significantly higher scores compared to people who had *not* had those experiences. These results, as illustrated in table 4.6, show that in all 11 dimensions there was a significant relationship between the semantic differential scale self-assessment scores and documented previous experiences. The adjusted R squared scores give an estimate (which is thought of as being conservative) of the overall influence the experiences themselves had on the scores. This varies from just under 5% to almost 20% in dimension (E). Although this influence is not calculated to be particularly strong, it does support the approach that was taken to strengthen the self-assessment questionnaire with the 'experiences' log.

For each of the 11 dimensions the influence of each specific experience was determined, and where there was a significant influence this was noted by a grey square in table 4.7. Whilst one cannot make sweeping conclusions about the significance of having a particular number of specific experiences influencing a dimension, it does give an indication of which experiences have the most potential to lead to the development of certain leadership capabilities. The importance of gaining a grounding in leadership learning from university / medical school and

the value of role-modelling was highlighted on more than one occasion. In addition work-based learning through experiences with patients, teaching and supervising colleagues and involvement in 'improvement' projects were important influences on a number of dimensions. This provides important backing for some of the work-based learning concepts that emerged from the literature review in chapter 2.

Conversely, more traditional methods of knowledge-focused 'classroom' learning such as lectures, teaching sessions, scenario-based learning and reading books / journals were perceived to be less influential experiences. The finding that participation in clinical audit was not found to be a significant influence in any of the dimensions leads to the suggestion that trainees have become much broader and more diverse in their approach to leadership learning. This pushes up against the fact that many postgraduate training programmes in the UK still use compulsory participation in a clinical audit in every 6-month long placement as the single way to involve trainees in the broader aspects of patient care. It is worth noting that there were some, perhaps slightly surprising, areas where some of the leadership dimensions did not significantly correlate to experiences that one might have intuitively thought would have had an impact on that learning. An example of this might be the 'managing yourself' dimension which was only significantly influenced by participation in improvement projects and teaching and supervising colleagues. That said, as illustrated in table 4.6, there was an overall influence of previous experiences on the managing yourself dimension, and with the relatively small numbers in the cohort, it is perhaps not unexpected that some areas of correlation will not be strong enough to be statistically significant.

The conclusions from this work on experiences have potential important use for postgraduate trainees and their trainers, as when looking to develop learning plans to address specific leadership learning needs, this information could lead to the delivery of slightly more targeted advice.

Chapter 5 – Qualitative analysis of the starting point

Chapter Overview

This chapter builds on the quantitative assessment of the previous chapter. It introduces the principles of framework analysis as a qualitative research methodology. It continues by describing the focus groups that were used to explore participants' ideas around leadership, and to gain a deeper, qualitative understanding of the baseline leadership learning of the cohort of participants, prior to the implementation of the intervention. The addition of this qualitative element is a key part of the mixed-methods approach. The chapter details the thematic analysis of this work and explores the conclusions that can be drawn from it.

5.1 Introduction

The previous chapter explored the quantitative self-assessment tool that was used to measure the pre-intervention starting point. This chapter builds on this to explore some basic theoretical concepts around the use of qualitative data in educational research, before describing the specific use of focus groups in this study. The aim of using these qualitative methods was to gain a deeper, more comprehensive understanding of the starting point (pre-test) of the cohort of participants. In addition this gave an opportunity to capture their ideas and learning needs and then to use these to shape the intervention accordingly.

5.1.1 Introduction to focus groups

The choice of focus groups (where a small number of participants come together in person to discuss their views on a particular topic) as a research methodology arose from the need to get underneath the straightforward quantitative answers from the self-assessment survey. The aim was to explore in greater depth the ideas that the paediatricians involved in this initiative had expressed. The value of focus groups in understanding the *why* questions was eloquently expressed by David Morgan twenty-five years ago (191):

> Focus groups are useful when it comes to investigating *what* participants think, but they excel at uncovering *why* participants think as they do.

Whilst there would have been a significant logistical advantage in either extending the self-assessment questionnaire to incorporate questions in this area, or designing and sending out a separate questionnaire to explore these issues, this was considered to be an inferior methodology in this situation. Had this option been taken, the likelihood was that the answers would have been more superficial and the respondents' perspectives more difficult to genuinely understand.

In her review of qualitative research, Rosaline Barbour highlights the different arenas in which focus groups can be applied. These include their use as an exploratory method, in seeking the student perspective, in studying change, in accessing the hidden curriculum and in understanding problematic areas of clinical practice (192). In this project the emphasis was on exploring perceptions and previous experiences, whilst strongly capturing the participants' voice, in order to then incorporate these ideas back into the intervention.

A crucial feature of a focus group interview is the group dynamics that underpin the discussions. If successfully captured, this means that the range and type of qualitative data generated through the social interaction of the group can be particularly deep and rich. The sections below describe the objectives and the data collection and analytical methods used, and the discussion section picks up on some of the rationale, advantages and disadvantages of this approach.

5.2 Objectives

The main objectives of this chapter, which link back to the overall project objective (d), and relate to the running of focus groups prior to the introduction of the intervention, were to:

- Explore the participants' perceptions as to what makes a doctor a good leader
- Consider which potential learning opportunities participants felt might help them to develop as leaders
- Relate these findings back to the quantitative data of the trainee committee cohort obtained from the self-assessment questionnaire described in chapter 4

5.3 Methods

This section describes the focus group methodology and thematic analysis used in the study. The methodology of the data collection and the analysis were chosen specifically to meet the

objectives described above. Each of the steps taken, and the rationale behind them are summarised in the sub-sections below.

5.3.1 Data collection

Each of the focus groups was run by the same pair of research assistants who had agreed a facilitation strategy with the lead researcher (RK) beforehand. One research assistant moderated the discussions (RH) whilst the other acted as a note taker (AR). Their role was to observe group dynamics and non-verbal interactions as well as noting down which statement was made by which individual. A digital audio recording was made for subsequent transcription and then analysis. As discussed in Chapter 1, participants were given an information sheet (see Appendix 5.1) and were consented for their involvement and use of the data produced (see Appendix 5.2).

Participants for the focus groups came from the trainee committee cohort, who were invited by e-mail to contribute on one of two different days. Three pre-intervention focus groups were run. They had 8, 7, and 6 participants respectively and lasted for between 35-45 minutes. Clinical duties for some trainees made it impossible to get all participants together at the same time so the first two focus groups were run on the same day (during the Trainee Committee induction day), whilst the third was run four weeks later. Both the timing of these focus groups, and their content, would justify them being labelled as part of the intervention. As noted in chapters 3 and 4, this 'testing' factor is a potential jeopardizing issue. Of note, the self-assessment web-based questionnaire was sent out to participants in the cohort group, and the comparator group, concurrently with the third focus group. This is relevant to discussions around 'sequencing' which are picked up in section 5.6 below.

Each of the focus groups was facilitated around the following three questions, which were designed by the lead researcher (RK):

- What makes a doctor a good leader?
- What opportunities would help you to develop as a leader?
- Thinking more broadly....what opportunities could the School of Paediatrics provide to develop (all) trainees as leaders?

The audio recordings were professionally transcribed into a digital typed document by an educational transcribing service. They were then independently checked by each of the three researchers (RK, AR and RH) against contemporaneous notes made by the assistant researchers (RH and AR) during the focus groups. The main transcription error that occurred was in the incorrect attribution of quotes to particular speakers. This was corrected as soon as the transcripts were returned, by each of the researchers independently re-listening to the digital audio recordings.

5.3.2 Data analysis

For this study, the methodology followed to analyse the focus group data was 'framework analysis', which was described by Krueger in 1994 (193), and was been chosen primarily because of the thorough step-wise approach that it entails. These clear, sequential steps are helpful when managing the complexity of large amounts of data.

> Framework analysis is an analytical process which involves a number of distinct, though highly interconnected stages (194).

The five key stages within the process of 'framework analysis' are: familiarization; identifying a thematic framework; indexing; charting; mapping and interpretation. Another important feature of framework analysis is that although it takes a structured thematic approach, it allows themes to emerge, both from the research questions asked, and from the personal narratives of each of the focus group participants.

5.3.3 Familiarisation & emergence of themes

Each of the three researchers (RK, AR and RH) familiarised themselves with the data by listening to the audio recordings and reading the full typed transcripts several times. The aim of this stage is to immerse oneself in the details of the data, whilst retaining the overall sense of the discussions, such that major themes start to emerge (193). It became clear from this process that by the third focus group very few new themes were emerging. The literature suggests (193)(195) that with fairly homogeneous groups only 3 or 4 focus groups are usually necessary to reach this stage, often called 'theoretical saturation'. This describes the point where the rate of appearance of new ideas and themes has slowed down to almost nil. Whilst it is possible to keep repeating focus groups with new participants (although in this study this would have been restricted by a fixed number in the cohort group), methodologically, when 'theoretical saturation' is reached this is an appropriate time to stop.

5.3.4 Identifying a thematic framework

This stage began by each of the three researchers individually annotating their versions of the transcripts with comments, phrases and ideas that emerge from the data. As this process continued throughout the full duration of each of the focus groups, key categories and headings emerged to form a numerically coded index. The three researchers then worked through the transcripts together comparing notes, categories and headings to reach consensus on the thematic framework that had emerged.

5.3.5 Indexing

The next stage, called indexing, involved sifting through the data, to highlight and sort out the quotes. This was done by systematically reading through the transcripts and applying the index to each of the individual quotes so that they could then be categorised alongside those which illustrated similar ideas and arguments – these are usually referred to as 'themes'. Following this the transcripts were annotated with numerical codes from the index, alongside brief descriptors to help expand on the index heading. Comparisons between different viewpoints could then be made. Single passages of text at times contained a number of different themes; where this occurred these quotes were recorded separately against each theme.

5.3.6 Charting

Once the quotes taken from the transcripts had been indexed they could be moved from their original context and placed into the thematic framework. This process, which used the 'cut and paste' editing functions of Microsoft Word, entailed organisation of the data as well as the combination of similar quotes to achieve data reduction. In addition to rearranging and grouping quotes, this process of charting also incorporated a significant amount of synthesis of ideas. The outcome at the end of this stage is that the charts contained distilled summaries of views and experiences of the participants as well as quotes of verbatim text.

5.3.7 Mapping & interpretation

This final stage of the framework analysis involves considering the relationship between the different quotes and ensuring that the data is linked together to provide a coherent 'big picture'. This is about using the data more broadly to define concepts, map the range and nature of the phenomena that have emerged, and find associations between themes, with a view to providing explanations for the findings (196). Krueger's original guidance (193) suggests the interpretation of data should include consideration of the actual words used, their context, their frequency and whether they are used by a number of participants, their internal

consistency, the intensity to which they were used, whether they referred to specific personal experience and whether they were part of a 'big picture'. More recent papers, including an update from Krueger, suggest a similar approach, but where some of the detail from the original guidance is excluded (197)(198), leaving frequency, specificity, emotions, extensiveness and 'big picture' as the five aspects of the data to consider. This was the approach taken in this work.

In this study, the process of mapping and interpretation was conducted by the lead researcher (RK) and then was subsequently reviewed by the other members of the research team (AR and RH). Any further comments and reflections from these discussions were built into the final version of the analysis, which is presented in the results section below.

Findings were integrated with some of the major concepts derived from the leadership literature review in chapter 1, from learning from the medical leadership development interventions described in the systematic review in chapter 2 and with the quantitative data described in chapter 4. Three different approaches to integrating data, including the use of a triangulation protocol and a mixed methods matrix were considered (67)(191). In the end the 'following a thread' technique, described by Moran-Ellis and colleagues, was adopted (199). In essence this is where a question or theme is selected from one dataset and then followed across another dataset – this is the 'thread'. This has the advantage of being fairly intuitive in its approach, but as with all integration methods a degree of expertise is needed to ensure consistency of approach (67).

5.4 Results

This section describes the results and analysis of the focus group work. As described in the objectives section, and expanded on in the methods section above, the focus groups were designed to:

- Explore the participants' perceptions to generate a hypothesis as to what makes a doctor a good leader
- Consider which potential learning opportunities participants felt might help them to develop as leaders
- Triangulate the trainee committee cohort's quantitative data obtained from the self-assessment survey

The 21 participants who were available to contribute to this part of the study were split into three focus groups as detailed in table 5.1. A count of the number of quotes from each participant that were adopted into the thematic framework was also made. Although a simplistic marker, this does give a broad indication of the relative contribution of each of the participants across the three focus groups.

Group A Participant Code	*Group A Number of Quotes*	Group B Participant Code	*Group B Number of Quotes*	Group C Participant Code	*Group C Number of Quotes*
3	*9*	4	*9*	1	*2*
8	*5*	9	*8*	6	*2*
12	*4*	10	*15*	14	*8*
13	*6*	11	*4*	21	*4*
17	*1*	15	*9*	27	*12*
19	*7*	18	*13*	28	*6*
24	*10*	20	*11*		
30	*6*				

Table 5.1 Allocation of pre-intervention focus group participants with illustration of the number of each participant's quotes used in the thematic analysis

The process of the lead researcher (RK) and the two junior researchers (RH and AR) working through the transcripts of the three focus groups and comparing notes, categories and headings led to consensus on the thematic framework that had emerged. Four broad themes (sometimes called 'categories' within the framework nomenclature) emerged from this framework analysis. These were:

1. The importance of leadership qualities in doctors as leaders
2. The importance of leadership behaviours in doctors as leaders
3. Specific opportunities that would help the participants develop as leaders
4. Considerations about the broader aspects of leadership training and learning

Underpinning these four themes, two underpinning dimensions emerged, which were also demonstrated within the factor analysis described in Chapter 4. The first of these dimensions was around the personal aspects of leadership development, including developing insight, managing oneself, acting with integrity and ongoing personal development. The second dimension was where participants described leadership behaviours and experiences that had a much more systems-based approach; these were around developing networks, understanding contexts for change and encouraging improvement and innovation. These dimensions fit

closely to the domains and elements of the Medical Leadership Competency Framework, and also reflect back strongly to the theories and concepts around transformational (22) and shared leadership (30) presented in chapter 1.

The four themes were 'evidenced' by the processes of indexing, charting and mapping described in section 5.3. This exercise mapped the participants' quotes to each theme, where together they provide cohesive emphasis. Quotes are attributed to individuals within each focus group and labelled with a 2-part code (eg GA; 24 or GB; 20) where the first two letters indicate which of the 3 focus groups they came from, and the number that follows being the individual's unique participant code.

5.4.1 Leadership qualities in doctors as leaders

Participants in the focus groups were quick to explain and discuss their views as to what qualities make a doctor a good leader. It is of interest that this links back strongly to some of the principles developed within leadership theories such as the trait theory, as described in sub-section 1.4.2. From the focus groups there was a strong emphasis on the importance of a good leader being inspirational to those around them:

GB; 10	"someone who can inspire other people"
GB; 9	"I am extremely happy working for this person and I would do anything they ask at a moment's notice"
GC; 27	"if you're an inspirational role model for people then they would be happy to agree with your decisions and follow you as a leader"

Personal values such as honesty, integrity and humility were also perceived to be crucial qualities of a leader:

GB; 10	"you would recognise them as very principled or, what's the word, with a lot of integrity and values"
GA; 24	"they're just exceptional leaders because they're very truthful"
GB; 9	"humility rather than arrogance is absolutely vital."

There were also interesting discussions about credibility, and exploration of what this really means and why it might be an important factor for doctors as leaders. The value of being a good, experienced clinician shone through in this argument; participants felt that as a leader you were more likely to be supported by colleagues if their perception of you was of someone with strong clinical credibility:

GA; 19	"when you think …if I was sick, or my child was sick, I would want to be looked after by them"
GA; 19	"respecting their clinical judgement"
GA; 8	"to be seen to be hard working themselves"

GC; 27 "people tend to look towards people who have got experience as well and will naturally engender respect through having been there and done it"

Participants also described the selfless motivation and self sacrifice that they saw in successful leaders:

GB; 9 great leaders as "the ones who really do put themselves out"

GC; 27 "it's often the people who are just there to care about other people and want to make things better that put the effort into doing those sort of things..... rather than the emphasis being to further their career"

GB; 9 "a degree of sort of self sacrifice, I don't know you could compare it to sort of leading your troops out into battle and being the one at the front who puts themselves most at risk"

They acknowledged that this was quite daunting, and reflected that this was a very challenging area for individuals interested in developing themselves for the future:

GB; 9 "I am not sure I really want to be doing this for the rest of my life."

Personal drive and ambition to take on leadership roles as a doctor were also important concepts that came up in discussion. There was debate as to whether the role of a leader was something only some doctors willingly took on, or whether it was something that was inherently part of being a Consultant.

GB; 4 "the desire to want to be in a leadership role... it is a matter of wanting to put yourself ahead as well which is something that doctors tend to do"

GC; 28 "a genuine desire to be a leader, as a doctor"

GB; 9 "I think in the role of Consultant….you take the flack for the things you make the decisions for and it ultimately rests on your head, and you run a team. It is clear that some people don't want to do that but that is what you have to do. That is what the inexorable path leads Consultants to."

Vision was another important quality that participants felt as being integral to a leader; this was both vision around the wider context and bigger picture of healthcare, as well as a personal vision that was constantly looking forward in a developmental fashion.

GB; 4 "somebody out there has to have the bigger picture in mind and a goal or objective"

GB; 20 "able to have innovative ideas and practical solutions to problems"

GB; 11 "they have this overall management plan which sort of takes you into account"

GC; 10 "[he] really looks ahead and moves the whole career forward"

The importance of courage as a leadership quality emerged from the focus groups, with particular emphasis on not being afraid to make difficult and unpopular decisions, and on giving honest and open feedback.

GA; 24	"they are very truthful; they might talk about things that other people wouldn't talk about and bring things up that perhaps everybody thinks about but won't mention"
GB; 18	"being a good leader also comprises being able to make the difficult decision or the unpopular decision that nobody else wants to make… you have got to have the humility to say 'this is a difficult situation… if nobody has a view, we are going to have to come down on one side or the other"
GC; 27	"it takes a very brave registrar to do that [give difficult feedback] and they've got to, I think, develop the experience of knowing their SHOs"

In summary, participants described a wide range of leadership qualities that they felt were important for doctors as leaders. Leaders needed to be inspirational, with drive, vision and courage to make the difficult decisions, but at the same time had to be credible, humble and selfless in their actions (6). The recognition of these values from within the trainee group emphasises perhaps the most important feature of all. Leaders are role-models, and through their behaviours and the example they set, the values described above are mirrored and copied by individuals within, and even beyond, their teams.

5.4.2 Leadership behaviours in doctors as leaders

Building on the discussions around important leadership values, the trainees in the focus groups explored a number of different behaviours that they believed were integral to successfully being a clinical leader. One of the most prominent concepts was around working and learning together in teams. It is interesting to reflect on how this fits in with some of the leadership frameworks that introduce ideas of shared (30) and distributed (28) leadership, and of followership (26). The trainees demonstrated the importance of understanding the perspective and values of others, embracing diversity and nurturing and supporting everyone within their team:

GB; 20	a great leader is "able to bring out the best in the team and making them work together well and realise that it does not matter if they have different strengths and powers, it matters if they can work together to make the team better."
GA; 30	"able to recognise people's strengths and weaknesses within the team"
GA; 3	"bringing the most out of people"
GB; 10	"[empowers] their trainees or their superiors to achieve their potential"
GC; 28	"being a bit of a mentor maybe, a motivator and a mentor.."

GB; 20 "making you feel, or their team feel that you are all in it together for a shared outcome"

An important aspect of the behaviours relating to successful team working is around being supportive. The trainees felt that this needed leaders to be trusting, easily approachable and readily available:

GA; 3 "giving advice and support when needed"

GB; 18 "I think being a good leader also is trusting the people you are working with to be able to do their part of the job"

GA; 12 "quick to respond and always approachable as well ....so someone can go to them"

GB; 4 "it makes a difference when they muck in.... willing to do all the crap as well as do all the stuff that they should be doing right at the top and it just makes you identify with them"

Trainees also felt it was important that a clinical leader was able to motivate their team to achieve success. The emphasis on these behaviours was around being able to take a facilitative, enabling approach that hoped to get the very best out of people:

GA; 8 "someone who can make their team actually perform to complete the objective that they are there for"

GB; 20 "they make you feel that all of you are working towards the same goal"

GC; 28 "what I think makes a good leader are those people that can motivate other people to either do similar things... a motivator"

GA; 3 "someone who can bring out the best in their team"

GB; 10 "I think someone who can inspire other people, empower I suppose, other people whether they are colleagues, their trainees or their superiors, to achieve their potential"

In order to be able to motivate and support others, leaders need to be outstanding communicators:

GA; 24 "being able to communicate to every member of the team effectively"

GB; 18 "somebody who can communicate that to the rest of the people they are working with"

GC; 21 "I think as a doctor, you're generally used to talking to other people.. so you should ideally have good communication skills which would be essential in leadership"

Participants in the focus groups recognised that effective communication as a leader is not just about issuing orders and commands, but is underpinned by having a facilitative, open approach where the views of others, wherever they are in the hierarchy of the system, are valued:

GB; 20 "listening and making people feel that their voices are heard but also making everybody feel that they have a good contribution to make even if it is quite different"

GA; 13 "for a leader to seek the opinions from anybody else they're working with"

GB; 18 "someone who is able to listen to the other people they are working with and make them feel that they have been part of the decision-making process"

GC; 10 "listen to what other people have to say and to respect the opinions and views of other people"

GC; 14 "being able to actually take in that information from various parties to be able to make a final decision"

GC; 21 "he always listened to the nurses' point of view and takes it in and he's also very good at listening to parents and their ideas and their expectations and so he commands a lot of respect through that which makes him a good leader"

An extension of these ideas around communication arose through discussions about the importance of leaders valuing their team, being kind and courteous and being openly appreciative of the efforts of others:

GA; 30 "I have always been impressed by bosses that have had a really horrible day at work and been very busy..... *and* said everyone's worked really hard today and that's really important"

GA; 3 "Just saying thank you is a huge thing that doesn't happen enough"

GB; 20 "making them [the team] feel very important and feel very valued"

GB; 4 "I think being valued and having been part of a team"

GC; 10 "he was still bothered about me as an individual and my career and my problems and I thought that was an excellent thing in a leader"

The trainees felt that an important mechanism for leaders to enable their teams to feel valued and appreciated was through promoting an atmosphere, or even culture, of learning and development within their teams.

GB; 18 "the good leaders I have worked for know when to just hang back; they are not unsupportive but they know when to hang back and let you try to do maybe more than you were expecting them to let you do"

GA; 8 "they recognise that you're doing quite well and either let you carry on or let you take that extra bit of responsibility"

GB; 15 leaders "create opportunities for development"

GB; 18 "you find at the end of time working for them, they have pushed you to actually do what you are able to do that you didn't realise you could do"

In addition to the different behaviours that allow a team to thrive and perform, there are a number of other leadership behaviours that were considered to be important during the focus group discussions. Perhaps unsurprisingly, the trainees recognised the value of experience, and the confidence that can come with this, as being important:

GB; 4	"something if you have been doing it for years… you feel you have the confidence to take the lead role"
GC; 27	"people tend to look towards people who have got experience as well and will naturally engender respect through having been there and done it"
GB; 11	"confidence. Somebody who knows what they are doing…"
GA; 19	"seem able to prioritise and keep everybody else calm... keep it all under control"

Making calm, transparent decisions was also recognised as being a vital leadership behaviour. Discussions recognised that clear decision making had to be underpinned by the ability to weigh up evidence, include the views of others and to be transparent in thought processes. Underpinning all of this is an important sense of responsibility:

GA; 13	"the ability to be able to explain the different pros and cons and different options"
GA; 8	"it's important the rest of the team can see the thought making process and how that decision was arrived at"
GA; 12	"communicating their ideas and thought processes"
GB; 9	"I think if you are working with somebody who can make a decision, and then explain to you exactly why they have done that, in both a teaching role and a leadership role, then that inspires confidence"
GC; 10	"I think a lot of doctors have to take charge of a situation...they are used to being authoritative and decisive which can be important qualities to have in a leader at times"

As highlighted in the introductory paragraphs of this results section, the trainees who contributed to the focus groups recognised that whilst one dimension around leadership development was around developing as an individual, there is a second dimension that is much more about the wider system. This understanding has a potential impact a leader's ability to successfully effect change:

GB; 15	"you know when they do make a decision they can follow it through and actually have got the aptitude to turn something either into something that will work… or actually create another opportunity"
GB; 20	"be able to have innovative ideas and practical solutions to problems."
GB; 11	"when you contact them or ask them something, you know it will get done."
GB; 15	"some of the most effective leaders I know understand the key people in other bodies really well"

In summary, participants described a wide range of leadership behaviours that they felt were important for doctors as leaders. Leaders need to recognise the importance of teams. They needed to be supportive and nurturing of colleagues, using facilitative approaches to get the very best out of people. They need to be able to motivate their teams, and to achieve this through open and enabling communication. Leaders need to have experience and confidence in their abilities so that they can make clear decisions. In order to be able to successfully lead their team through change, they need to be able to contextualise this decision-making within the wider system in which they work. Crucially they need to be appreciative of the efforts of their colleagues, value the opinions and ideas of other people and create a culture of learning within the workplace.

5.4.3 Opportunities to support leadership development in the workplace

Having heard the trainees' views on the values and behaviours that they perceive as being integral to clinical leadership, the discussions within the focus groups moved on to explore their opinions on how they thought they themselves might develop as leaders. The aim was for them to consider different ways and opportunities through which they might learn and develop. Much of the discussion was around the development of appropriate knowledge and behaviours alongside all of the clinical learning that takes place in the workplace.

Interestingly, these discussions were contextualised within a series of comments underpinned by anxieties around self-awareness and that feeling about stepping up to Consultant level that 'I will never be ready'. This makes a strong case for the value of learning from role-modelling, and the importance of supporting doctors to develop a strong ethos of reflective practice from the very beginning of their careers:

GB; 10	"some people feel 'I don't want to be a Consultant because I won't be a good leader' and actually they might just be limiting themselves by their own perception of what it is"
GB; 18	"need to make them think about the things that may be innate to them, that maybe they don't appreciate in themselves, that they could develop over time...these will help them to fulfil the role when they reach it"
GC; 27	"when I watch a consultant I try and see what's good and what's bad, what would I want to pick up and what do I not want to do"
GA; 30	"so hopefully, if you take the good things from good leaders that you see, that will make you a good leader as well"

Many of the trainees were clear that the levels of understanding as to what 'leadership' means (and also does not mean) were generally poor amongst doctors, and that this was a fairly prohibitive factor in achieving a strong leadership development culture across all healthcare

institutions. The demands of all of the clinical learning, with no time of focus on anything else, can often mean that there is no space or priority given to leadership learning. This is often reinforced by the attitudes of clinical and educational supervisors who demand that the sole focus of trainees is on their clinical learning. Raising awareness of what leadership actually means, and making the link to high quality patient care, were emphasised as important strategies going forward:

GA; 3	"this is about improving reputation of leadership away from those awful away days and stuff like that"
GB; 10	"you don't want [leadership and management learning] to be an imposed sort of chore… if you can foster it on a small scale early then people can have a positive experience of it and actually believe in the process but if you kind of say 'everyone has to do management and it's going to be defined like this and everyone has to do this', then it is just another thing to add to the millions of other things that we *have* to do."
GA; 19	"I think making it clear that leadership is really important because it's quite easy to get sidelined with all the other things…"
GC; 1	"they should start really early on… that it becomes such a fabric of what you do as part of medicine."
GC; 10	"I'll bet there are things that I will think about now which would be useful ... which are probably obvious once you've heard them."

There were also a number of discussions around what people thought might be the best timing of leadership development for trainees. Historically, this has almost always been learning that has only begun in the weeks before a consultant interview; discussion around compulsory versus voluntary, course-based versus work-based had proponents on both sides of the argument. There was clarity that in order to be successful trainees need some space and exposure to these ideas in addition to their clinical learning

GB; 11	"I think key is giving people exposure to what leadership is."
GB; 4	"to recognise the time put aside, the sacrifice of time for this, you have to have recognition from up on top saying 'yes we will carry the can for a short while and let our juniors go away, we will do the work whilst they go off and do this, because it is worth doing because it will be of benefit to us in the future."
GB; 9	"you either have to have an element of compulsion I think in order to get people to do it or you wait a bit later to when people are worried about not having the leadership skills they are much more imminent about needing."
GB; 18	"it does need to be something that is not sidelined as a voluntary 'oh well I'll do it when I've got time and I can get study leave'"
GC; 21	"there was a definite difference looking back now about how I acted with my SHOs in the first year [as a registrar] and how I did with the second year as a reg and I think it's just something that you perhaps can't teach and just comes with experience"

There was also recognition that doctors needed to be challenged and given leadership roles, but in a graded and supported way so that they did not feel immediately out of their depth.

GA; 30	"we need to be able to be given that leadership role, even if we're not quite ready for it at that stage and that leaders actually need to take a step back and watch you do the leading things"
GA; 13	"it's usually through something else in addition to your usual stuff that you meet other people who are leaders"
GC; 27	"giving people at particular levels graded opportunities to get used to what essentially I think are leadership roles. I think giving somebody the chance to be in charge of the rota is different from dropping someone in it."

The discussions that explored how deep this learning should be were also very varied. It seemed that the main concept to come out of these discussions was that a degree of leadership learning was necessary for everyone, but that greater depth of leadership development may only become relevant if taking on specific leadership roles in the future:

GA; 24	"it's not about everybody being a great leader, and that they want to become a chief executive ....it is just the sense that you can be the leader in your department or just your group, or your team ..you need to know how to go about making changes in services for the benefit of patients"
GB; 4	"not necessarily about teaching you to be an absolutely fantastic leader but it's just about the strategies for coping and improving your basic management… so everybody who pops out of the system has the competencies to be a competent manager or a competent leader."
GC; 28	"I'm almost beginning to think that there are two levels of leadership here, that there's a baseline level that all doctors need to know to cope with their working day, and then there's a good leader that you want in charge."

In terms of where the biggest gaps in learning and prior experience were, the trainees generally felt least knowledgeable around many of the concepts relating to systems and the wider context of healthcare. This correlates well with the findings from the quantitative work discussed in Chapter 4:

GA; 24	"alongside all of this would have to be some training about the system we work in… how the NHS functions, how the Trust functions, what a Trust board is"
GA; 12	"I think it's important to know about that [understanding Trust structures and decision making process] so you respect your leader and their decisions …as it affects your clinical work as well".
GA; 24	"all of us, if we understood the way that our services worked, could make a small improvement and that in itself is leadership quality."
GC; 27	"we don't get a real sense of their strategic input into services... often you're quite hidden away from those management aspects"

Interestingly, there was strong recognition of the value of looking outside paediatrics, and even outside healthcare, for insight and guidance around how one best organise leadership development:

GB; 4 "if you look at the corporate sector, if they are training anyone in management or a higher management role….they invest huge amounts of time and money into doing that."

GC; 14 "the reason why these big corporate, big companies pay so much money for them [corporate away days] is because it is value,"

GA; 13 "to think about leadership, not just in medicine, but also the links with other types of business…."

GC; 14 "GPs as we've heard do run the whole system a little bit more like a business and obviously from that point of view they are slightly further forward looking than most other specialties ...and I think we should try to learn from these things."

The trainees were clear that in order to develop clinical leadership competencies they needed sufficient space and independence to learn and reflect; they highlighted a number of different barriers to this learning within the current system, and also put forward some potential solutions:

GB; 20 "I think however motivated and intelligent people are, sometimes they just don't find the time to devote to everything that they want..... if they are petrified at the thought of a child fitting in resus and they don't know what to do, they might feel that they want to spend their time developing those clinical skills first."

GA; 12 "they've also got to find time to take a day off for annual leave, study leave… to go and do a leadership course"

GB; 20 "they just feel that they can't or don't have the brain space for it.

GB; 4 "it is very hard for people, particularly in the NHS when you have got a strained work environment, to take people off the on-take system, off the ward, to go and do this. But I think that is the only way it will actually happen."

GC; 27 "I think a good leader is someone who will look for that responsibility and then look to reflect on how they do it when they take it"

Acknowledging that dedicated time out of clinical work is difficult to achieve, certainly for the majority, the focus of the discussion moved to considering maximising on-the-job leadership learning. Ideas around shadowing, and learning from role-modelling were strongly represented:

GA; 13 "you learn from bad leaders. You watch them do things badly and you think I'm never going to do it that way"

GB; 15 "a placement or taster where you work alongside a more senior colleague in a different hospital because you have heard that they are particularly good at 'X'"

GB; 15 "there are some consultants they think are better role models than others."

GC; 27 "often Consultants lead outside of the clinical sphere but we don't get a real sense of their strategic input into services... often you're quite hidden away from those management aspects"

GC; 14 "we also shouldn't forget the importance of learning by example, I think a lot of us will probably become slightly better leaders because we have seen people in leadership roles do them well and but actually the other way around as well, people who have not done them so well and learnt from their mistakes"

GC; 27 "when I watch a consultant I try and see what's good and what's bad, what would I want to pick up and what do I not want to do"

In addition to the more passive aspects of watching consultants' behaviours and looking up at role-models, there was enthusiasm for a more active model of leadership learning through coaching and mentoring:

GB; 15 "you sort of pick your own supervisors that you hang in contact with longer and you can work under them and develop your own projects."

GC; 28 "mentorship, that's the cornerstone of it all, I think people need to be mentored into being good leaders"

GA; 24 "I think if all of us recognised that actually if you approach people and give them an idea in a certain way, most people are responsive to it"

GB; 15 "I think that we can also learn from different institutions that we are at."

Feedback was another well-discussed area within the focus groups. The trainees highlighted that that vast majority of opportunities for feedback are missed, and that addressing this would be significant progress. They talked about the need for clinicians to learn to give, seek out and accept meaningful and developmental feedback, and recognised some of the barriers to this around time, opportunity, hierarchies and above all culture:

GA; 6 "if someone feeds back.....when things maybe have gone a bit wrong, or you've had a very busy shift, and someone more senior... sits down to go through how we could make it better if you were in that situation again"

GA; 13 "so I suppose actually it ought to be a more formal training on how to give good feedback… how to give it in a more positive way…"

GC; 28 "Often you sit in committees and you may or may not have a say, but you still feel like you're just a junior…there's got to be some positive feedback within those systems to make you feel like you're a valued member and therefore that you will process"

GA; 3 "there isn't the pathway up to feedback up to other people leading you. I don't know, it still feels slightly difficult about that"

Perhaps unsurprisingly, amongst what were self-selected groups, the trainees put a strong onus on personal initiative and self-motivation in identifying and making the most of leadership learning opportunities:

GA; 19	"I think it's actually making an effort to reflect on something you've done, that could be quite a useful way as well of improving those skills"
GB; 20	"I think that if one wants to develop leadership skills and qualities, there are opportunities to do it, if you take them, within the job that you are doing"
GB; 10	"people want to have someone come to them with ideas… yet I think possibly as juniors, we don't feel that we can take the reins on certain issues so we just limit ourselves to traditional management duties"
GB; 15	"you have control over what you want to do for your future and you can take samples from different places with all your different experiences."

They recognised that for this type of leadership learning to become more widespread across postgraduate training and beyond, this culture of learning from opportunities and experiences that are going on in work on a weekly basis, would have to be developed at a more systematic level:

GA; 19	"it's something you should be trying to include in your work"
GA; 17	"that's definitely on the job training for leadership that I think I found very useful"
GA: 24	"it may be something that's going to be put into the curriculum at medical school"
GB; 18	"I think if you kind of chipped away at it, maybe even from the very beginning, through having management and leadership skills sort of integrated into the curriculum, the way we teach everything else, it doesn't then become about box ticking."
GB; 4	suggests "embedding learning in simulation"
GC; 14	"I think something like that should be properly built into the system as well"
GC; 1	"that it's really easy to put on a two day leadership course and make it part of a training programme but you've got to be careful that it's not just that. You might take something away from it but on it's own it probably doesn't really mean very much."

In order to achieve this it was highlighted that trainers, and those in supervisory positions, have very important roles to play in facilitating and enabling this learning to happen:

GA; 30	"in some ways, you're going to have to train the trainer first of all"
GA; 24	"it's really important to have an enthusiastic lead already in situ, in your workplace"
GC; 14	"Absolutely and once again that's something we can learn from GPs because GP trainers have allocated free time to actually do these things."

The trainees went on to explore ideas around how one might assess and ensure the quality of leadership learning. It was acknowledged that this is an important area, although there was concern that the more prescriptive nature of learning outcomes and assessment, with an over-focus on knowledge, rather than attitudes and behaviours, might actually stifle learning. The trainees thought that multi-source (360 degree) feedback was likely to be the most useful assessment tool in this area:

GA; 30	"you have to think about assessing it at some point to prove that that person has met the level"
GA; 3	"once it's been integrated, you've got to find a way of assessing it"
GA; 8	"you could have a section in 360 degree appraisal where people could assess how they perceive your leadership"
GA; 6	"There's a formal assessment on that 360 degree appraisal where everyone says they can communicate or they can't"
GB; 10	"I don't think you gain a lot from being too prescriptive. I think definitely exposure for sure, but not necessarily directed outcome measures. I don't think you need that. I don't think it adds value."
GB; 18	"I think it is something which is very difficult to assess and to say 'this makes you a good leader, this makes you a bad leader'. I think assessment would almost be counterproductive if you are talking about teaching people leadership skills."

In summary, having explored ideas around leadership qualities and behaviours, the trainees in the focus groups looked in depth at different opportunities that have given them leadership experience and learning. They acknowledged the importance of role-modelling, coaching and mentoring, both to help reflect on learning, but also in helping them to get over that feeling of 'I will never be ready....'. They recognised the importance of being really clear with potential learners as to what leadership actually means; this was not to be left to assumption. They advocated early exposure to leadership opportunities, but emphasised that this needed both support and space and should be introduced in a graduated way. They thought that there was a particular need to learn about and understand the wider healthcare system, and that rather than solely relying on the self-motivation of individuals a more embedded, systematic approach to leadership development should be established. Trainers, through highlighting opportunities and learning how to give honest, open, developmental feedback, would be crucial to achieving this. Whilst acknowledging that assessment might be an important lever for learning, they were concerned that if done too formally, it might suppress learning around the development of leadership behaviours.

5.4.4 The broader aspects of leadership training and learning

In addition to the exploration of ideas around leadership qualities, behaviours and opportunities for learning, there was a concluding category that emerged, looking at some of

the broader aspects of leadership training and learning. In particular these discussions considered who should be receiving leadership training. There was a predominant voice from the trainees that some sort of leadership role is inevitable for *all* doctors, and that therefore leadership development programmes should become an integral part of training for everyone:

GB; 18	"I think there is a certain inevitability to having to do some form of leadership, in whatever form that takes."
GC; 10	"as a registrar you have a position of leadership and as a consultant you're going to have a position of leadership.... everybody leads whether you're a natural leader or not"
GA; 19	"you can probably, certainly to an extent, learn leadership skills and I think that at the moment, we're not taught enough"
GA; 24	"You don't have to aspire to be some great leader… just all of us, if we understood the way that our services worked, could make a small improvement and that in itself is leadership quality."
GB; 20	"I think some people have natural leadership qualities that they obviously exhibit and then having formal training in leadership can maybe bring those to the surface."
GB; 18	"if you start early on with people and just sit down as part of a teaching thing in small groups and just say 'this is what leadership is, as a consultant and as a doctor; it is an inevitability - part of your job is a leadership role'. I don't think people necessarily think about it like that, but it is."

However, there were a few opinions cautioning against making everyone participate in leadership development activities, with suggestions of leaving this as an area where doctors could self-select themselves to get involved. Interestingly these goes somewhat against the direction of travel set by proponents of shared leadership approaches (30)(117), and that of the Medical Leadership Competency Framework itself (19). It was highlighted that for self-selection to be successful potential participants need to have good levels of self-awareness:

GC; 14	"I think the original premise that basically people who are more interested in doing certain things should be allowed to develop their interest in those fields rather than trying to make everybody do everything"
GC; 27	"Give them the route to identify how they can self-select themselves [for leadership training]"
GC; 21	"[to] go onto that true leadership role you have to be self-selecting to want to do it.....there's nothing worse than having someone that doesn't want to be a leader and it's forced upon them"
GB; 18	"there are certain people who do have good leadership skills but aren't aware of it… and if you make it a very voluntary process you are automatically going to select the people who think they have great leadership skills… and you maybe are going to miss those people who are a gem in a cupboard somewhere."

Overall, the feeling was that the key was to offer work-based learning opportunities from early on, and that in small steps, simple projects to improve a small aspect of care, and graduated levels of responsibility, were likely to make the biggest difference going forwards. This builds on some of the theories of experiential learning (109), learning from other leadership development interventions (100)(111) and on the findings from the experience log and self-assessment scores described in sub-section 4.4.5.

GA; 3	"I think it's at a very simple level where you need to give the opportunities to the juniors to make sure that they are given responsibility"
GB; 10	"when they then started their jobs, their first jobs, they just had that background understanding of what their seniors were taking on board and the decisions they made"
GC; 14	"early exposure to these issues, management type issues, leadership issues because… at a lower level we don't get exposed to these things, and often most of us don't really know what's happening behind the lines."
GB; 15	"learning the practicalities of how you actually do it on the job rather than learning leadership principles which are quite difficult to actually put into practice."

In summary, despite some caution from a few of the trainees, there was an overall feeling that leadership learning should be for everyone. Regardless of whether or not an individual had any aspirations to become a senior medical leader, all doctors need to demonstrate leadership capabilities within different aspects of their work. The trainees felt that this should start early in their careers and be built upon in small steps throughout their training.

5.5 Conclusions

This section describes the conclusions that can be drawn from the focus group analysis which was documented in the results section above. The analysis showed four broad themes to have emerged, all of which relate back to the main objective of the chapter which was to explore participants' ideas around leadership as well as gaining a deeper, qualitative understanding of the baseline leadership learning of the trainee committee cohort.

These four emergent themes were:

1. The importance of leadership qualities in doctors as leaders - leaders needed to be inspirational role-models, with drive, vision and courage to make the difficult decisions, but at the same time had to be credible, humble and selfless in their actions.
2. The importance of leadership behaviours in doctors as leaders - leaders needed to recognise the importance of teams, be supportive and nurturing of colleagues and use

facilitative approaches to get the very best out of people. They also needed to be able to make decisions, and to do this within the context of the wider system.

3. Specific opportunities that would help the participants develop as leaders – doctors needed to be challenged and given leadership roles, but in a graded and supported way so that they did not feel immediately out of their depth. They needed work-based opportunities for leadership learning, alongside the space to reflect on their experiences. Role-modelling, mentoring and coaching are crucial for this.
4. The value of broader aspects of leadership training and learning – underpinned by the thinking that all doctors should have opportunities to develop as leaders, and that this process should start early in their careers.

As outlined in the overall project objectives, in addition to providing triangulation to the baseline quantitative analysis outlined in Chapter 4, these four themes could be used to shape the content and learning outcomes of this and other leadership development interventions.

5.6 Discussion

The adoption of the mixed methods approach outlined in the latter part of chapter 3 is key to the exploration of the research question. At the same time as capturing some baseline quantitative data, as described in chapter 4, focus groups were used to gather more qualitative data from the trainee committee cohort. The decision to use focus groups, as opposed to relying solely on questionnaire derived data, was based on the need to both explore *what* participants were thinking and also *why* they had those views (191). There was also the advantage of using the relatively similar backgrounds of the participants (ie all had studied medicine, had chosen to be paediatricians, had applied to be involved in the trainee committee) to present ideas to each other for challenge and refinement (198) (197). The synergy gained from these group interactions, each with between 6 and 8 participants, is a well recognised benefit of focus groups (200).

Interviews are another important method for capturing this type of qualitative data and are particularly strong in terms of exploring why individual participants hold specific views. They hold some advantages over focus groups in terms of being able to achieve greater depth of conversation, be more targeted, and give more time to each individual to explain their views (201). There are also advantages that interviews can reduce the possibilities of cross-contamination of data, and avoid the risk where one person dominates a focus group (192). This one-to-one environment can also be more conducive to discussion of issues that are

sensitive or personal. However, in this study, due to the nature of the topic being explored, it was not expected that there would have been difficult discussions about personal issues that were then not surfaced due to the nature of the group.

While these potential gains from interviewing were considered, it was decided, on balance, that firstly there was not sufficient resource of time for the lead researcher (RK) to conduct somewhere between 20 and 25 individual interviews. Secondly, the potential advantage of capturing data from synergistic discussions within each of the groups was felt likely to be particularly helpful in bringing depth to this qualitative work (191)(197).

Having made the decision to use focus groups to capture data, a method of analysis needed to be chosen. There is a huge amount of literature, particularly in the social sciences arena, where the use of different data analysis methodologies available to focus group researchers is hotly debated. The underlying principles are less controversial; in order to minimise the potential bias introduced in analysing and interpreting focus group data the analysis needs to be systematic, sequential, verifiable, and continuous (197). In reality, most researchers end up using some sort of locally adapted hybrid method of analysis that incorporates a combination of approaches (200). For this study, the methodology used was 'framework analysis' (193). This was chosen primarily because of the thorough step-wise approach that it entails. These clear, sequential steps, which are detailed in full in section 5.3, are helpful when managing the complexity of large amounts of data. This method also adopts an approach of continual data analysis, which occurs concurrently with data collection, as well as after it. It can be argued that with this method the analysis does not take place in a linear form and that one part of the process overlaps the next.

These processes of coding and retrieving described in the methods section of this chapter were very time consuming and laborious, and in many other circumstances are now being undertaken by software packages specifically designed for the purpose. In addition to searching for data some packages are underpinned by more complex algorithms which can identify co-occurring codes that may have been missed through other methods. However, there are potential problems with these packages and none are capable of making the link between theory and data, nor are they able to determine the structure of the analysis. Pope et al (196) recognise the value of software helping with the more laborious aspects of analysis, but suggest caution is needed:

> The prospect of computer assisted analysis may persuade researchers (or those who fund them) that they can manage much larger amounts of data and increase the apparent "power" of their study. However, qualitative studies are not designed to be

> representative in terms of statistical generalisability, and they may gain little from an expanded sample size except a more cumbersome dataset. The sample size should be directed by the research question and analytical requirements, such as data saturation, rather than by the available software.

Despite the potential availability of an analytical software package for use in this study, the importance of systematically working through and understanding each of the methodological steps in generating theory was highly valued.

Looking back on the systematic review described in chapter 2, there were only two included studies that attempted to combine pre- and post-intervention quantitative comparisons with qualitative data gained from semi-structured interviews (91) (94). None of the included studies used focus groups in this situation. The rationale of running focus groups at this stage of the study was to hear from the participants about their views and ideas and to understand some of the thinking behind them. These results could then be reviewed alongside the quantitative data in order to help make sense of them and to draw out any resonant themes. This process generated some interpretation of the quantitative results as well as being able to use the ideas and themes that arose from the focus groups iteratively to influence the design of the programme. In addition the focus groups allowed for a more 'narrative' documentation of where the different participants were in their thinking and their personal reflections of their own leadership learning.

As highlighted in the methods section above, the first two focus groups were held shortly before the self-assessment web-based survey was sent out, and the third focus group at the same time. Alongside the order in which the different components of mixed methods research are analysed and the way in which data is integrated, this 'sequencing' of the data collection is important to consider in analysis of the results (202)(203). As described in Creswell's comprehensive text (147), mixed methods research comes in many different shapes and sizes; these factors are key considerations when interpreting the data. This study attempted to use the concurrent triangulation approach. As detailed in sub-section 3.3.8, for both the pre-test and post-test evaluation, the aim was to collect data concurrently so that the two qualitative and quantitative data sets could be compared to see whether there were any differences, convergence or some combination of the two. Although this approach can lead to well-validated and substantiated findings, as noted in chapter 3, there are limitations to this methodology (147). It requires significant effort and expertise to study a single phenomenon with two separate methods, and it can be difficult to compare the results of the two analyses using data of different formats. It can also be difficult to work out how to deal with discrepancies in the data collected by the two methods.

Thinking about the sequencing in this study, the pragmatics of organising three focus groups for a defined cohort of participants meant that the first two focus groups were run 4 weeks before the third focus group and the issuing of the self-assessment survey. It is important to acknowledge, therefore, that the discussions within the focus groups may have influenced how individuals then went on to answer the web-based self-assessment survey when it was sent out to them. It could be hypothesised that their thinking may have been deeper, and more comprehensive than a peer who had not already participated in the focus groups. This is by no means a 'fatal' error in the methodological design, but needs to be considered when interpreting results. It is also important to acknowledge that, much like the discussion about the self-assessment questionnaire in chapter 4, the focus groups themselves should be considered as part of the leadership intervention, and in that role are a potentially jeopardising 'testing' factor. This is particularly important when interpreting the results and thinking about how the intervention might be re-run elsewhere.

The themes that arose from the qualitative analysis focused on the importance of leadership qualities and behaviours, with the trainees suggesting that the development of leadership competencies requires work-based space, time and opportunities for leadership learning. These ideas around needing the space and time within the workplace for leadership learning give some interpretation to the finding from chapter 4 that demonstrated a significant correlation between the number of previous experiences a participant had gained and their self-assessment score. Whilst the qualitative work highlighted a real breadth of different suggested learning opportunities the emphasis on the importance of support, mentoring, enabling and role-modelling from senior colleagues was consistent throughout the data. This correlates well with the conclusions of many of the medical leadership studies described in chapter 2, where the success of many of the leadership development programmes was underpinned by a culture of support and mentoring.

Although by this stage, the broad framework of the leadership development intervention was in place, these 'user-views' were able to be fed back into the final design of the intervention. I have acknowledged earlier in the thesis (chapter 3) that this is an approach that creates some contamination issues which might impact on the interpretation of the results. The positive side of this was the addition of some really helpful inputs from the trainees that enabled their leadership development programme to have a significant bespoke component. A specific example of this would be the organisation of workshops with focus and learning outcomes designed to address trainee anxieties about concepts around systems and the wider context of healthcare.

Having completed the ‘starting point’ (pre-test), or baseline, quantitative and qualitative measures the leadership development programme, described in detail in chapter 3, could be introduced. The following two chapters explore the impact of the leadership development intervention, through describing the evaluation methods, results and analysis from the end of the study.

Chapter 6 – Quantitative measurement of the end point

Chapter Overview

This chapter examines the quantitative methodologies that were used to measure the post-intervention (post-test) 'end point' of the study and hence evaluate the impact of the leadership development intervention. The chapter details the comparative results of the pre-intervention (pre-test) and post-intervention (post-test) self-assessment scores, describes the rationale behind the different statistical analyses used in these comparisons and explores the conclusions that can be drawn from them.

6.1 Introduction

The previous three chapters have described the development of the intervention in this study and the mixed methods evaluation strategy (chapter 3), the quantitative measurement of the starting point using a self-assessment tool (chapter 4) and the use of focus groups to gain a qualitative understanding of where participants felt they were prior to the introduction of the intervention (chapter 5). This chapter moves the study forward to 12 months later, and addresses the quantitative methodologies that were used to measure the post-intervention (post-test) 'end point' of the study. In doing so, through a comparison of the pre-intervention and post-intervention data, it aims to evaluate the impact of the leadership development intervention.

6.2 Objectives

The objectives of this chapter relate tightly back to the fifth overall project objective that is outlined in sub-section 1.3.2, namely to evaluate quantitatively the effectiveness of the leadership development intervention experienced by the trainee committee cohort.

6.3 Methods

This section outlines the quantitative methodologies used in the measuring of the post-intervention 'end point'. It details the design and implementation of the self-assessment tool, and describes and critiques the statistical methods used to analyse the results.

6.3.1 Use and implementation of self-assessment tool

Sub-section 4.3.1 within chapter 4 described in detail the development of the pre-intervention self-assessment tool. This included a description of how content validity was gained, the development of an 'ability scale' for use within the tool, and an introduction to a methodological approach to 'evidence' self-assessment scores. The methods and discussion sections of chapter 4 describe some of the opportunities and difficulties with the design and methodology used.

The post-intervention self-assessment tool was adapted directly from this pre-intervention self-assessment tool (see Appendix 4.1). The introduction was changed so that it acknowledged that participants had completed an identical survey 12 months previously, but it used the same set of 11 questions, each with two short vignette examples, the identical tick-box ability scale and the same section in which supporting experiences could be noted. This consistency was important in order to be able to make meaningful pre- and post-intervention comparisons, although the re-use of questions and experience in themselves may lead to some issues with internal validity (146).

This post-intervention self-assessment questionnaire was also designed and built on the www.surveymonkey.com web-based survey platform. 12 months after the pre-intervention survey had been distributed it was sent out to the trainee committee cohort via individual e-mails. In terms of considering the impact of sequencing, this happened concurrently to the first two post-test focus groups described in chapter 7. As previously, the technique where the main introductory message and survey URL link was pasted into the body of each e-mail to each individual potential participant, allowed for personalisation of the request to complete the survey.

A single, individual e-mail reminder was sent to potential participants between 14 and 21 days after the original request. The survey remained open for a further 8 weeks although the majority of the responses came within 72 hours of sending the requests.

6.3.2 Statistical analysis and rationale

As with the pre-intervention data described in chapter 4, the data from the web-based post-intervention self-assessment questionnaires was initially downloaded into Microsoft Excel format, before being manually cleaned in preparation for being converted to SPSS format. The analysis was performed within SPSS; as with the pre-intervention data support and guidance

for this, the choice and running of statistical tests arose from discussions with Henry Potts, an experienced medical statistician within University College London.

6.3.2.1 *Analysis of variance*

As with the initial tests to compare the self-assessment scores of the trainee committee cohort and the comparator group in chapter 4, the analysis of the post-intervention results should begin with a 'global test' to determine whether there was an overall effect of the intervention. Even if only a few, or even none of the 11 questions showed any *significant* difference as individual questions, there remains the hypothesis that there might have been a 'small amount of overall change over the 12 months of the leadership development intervention across the 11 dimensions', and collected together this might amount to a significant difference. This concept was analysed on SPSS using a global analysis of variance (ANOVA) test. Should this ANOVA test demonstrate a positive result then the next step would be to move on to post-hoc tests, such as paired t-tests, that could be used to examine each individual dimension.

6.3.2.2 *Paired t-tests*

Within sub-section 4.3.1.2 the debate as to whether the data from Likert scales is interval or not is outlined. As with the pre-intervention data, histograms of the new data were generated and reviewed. As previously, they looked to be normally distributed, with no evidence of a long tail (this would be very unlikely with only 5 data points) or skewing of the data. This meant that it was possible to use parametric statistical tests on this data, so the pre- and post-intervention scores of the trainee committee cohort for each individual dimension were analysed using a paired t-test. The 'paired' component of the t-test arises as it is a comparison of data points from the same group across two different points in time (ie pre-intervention and post-intervention). The rationale for chosing this parametric test is discussed in detail in chapter 4.

6.3.2.3 *Bonferroni correction*

As with the pre-intervention data analysis in chapter 4, the use of Bonferroni corrections (189) to account for the relatively small samples sizes was considered. This adjustment, which is widely viewed as being quite conservative, reduces the critical level of 'significance' from 0.05 by dividing it by the number of tests being run. With the 11 different questions within the self-assessment questionnaire, and the data arising from each, this would lead to a revised cut off for 'significance' of 0.05 divided by 11, which is 0.0045. This means that there would have to be fairly large differences in scores between the pre- and post-intervention results, to attain

significance. As explained in chapter 4 there is an argument that in the case of post-hoc tests, being run after obtaining a positive global test analysis, results can legitimately be displayed without applying the Bonferroni correction, and this is discussed further below.

6.4 Results

This section describes the quantitative results of the work to measure the 'end point', 12 months after the implementation of the leadership development intervention. The data described in this section were generated from the post-intervention self-assessment tool outlined above.

6.4.1 Response rates

The URL link to the self-assessment tool was sent by e-mail to the 30 of the 32 trainee committee members who were still involved and who had completed the full 12 months within the committee. One trainee had moved overseas for a research fellowship and another had moved out of London and so was no longer able to participate. 24 of these 30 trainees (80% response rate) fully completed the post-intervention self-assessment questionnaire, which compares with the 27 out of 32 (84% response rate) who fully completed the pre-intervention self-assessment questionnaire. However, the trainees who did not complete the pre-intervention questionnaire were not, with one exception, the same trainees who did not complete the post-intervention questionnaire. Therefore, in total, 22 out of the original cohort of 32 trainees fully completed the pre-intervention questionnaire, participated in the full 12 months of the leadership development programme and fully completed the post-intervention questionnaire. This gives an overall combined completion/response rate of 69%.

As explained in chapter 4 there was no post-intervention assessment of the comparator-group. This was not needed as the exploration of the differences between the scores of individuals of different levels of seniority (in sub-section 4.4.4) had shown that there was very little 'natural' increase in scores as the trainees went on through their years of training.

6.4.2 Comparison of pre- and post-intervention scores

Mean scores were calculated from the 11 self-assessed 'ability scales' on the pre- and post-intervention self-assessment questionnaires completed by the trainee committee cohort. The results are illustrated in Table 6.1 and detailed in full in Appendix 4.1.

It is important to note that in the post-intervention scores the means have been calculated from the number of participants who completed that particular question.

Dimensions	**TC cohort (pre-); mean** N = 27	**TC cohort (post-); mean** N = 24 unless stated
(A) Developing self awareness	3.5	3.5 (N=26)
(B) Managing yourself	3.7	3.7 (N=25)
(C) Continuing your personal development	3.7	3.8
(D) Acting with integrity and managing people	3.4	3.4
(E) Developing networks and building & maintaining relationships	3.3	3.9
(F) Working within teams and encouraging contribution	3.6	4.0
(G) Planning and applying knowledge & evidence	3.3	3.2
(H) Managing resources & performance	3.2	3.5
(I) Critically evaluating, ensuring patient safety and facilitating transformation	3.1	3.4
(J) Identifying the contexts for change and encouraging improvement & innovation	3.3	3.6
(K) Making decisions and evaluating their impact	3.2	3.3

Table 6.1 Comparison of pre-intervention and post-intervention self-assessment mean scores in the trainee committee group.

This demonstrates that the mean trainee committee cohort scores (displayed to one decimal point) did not change in time over the 12 months in 3 of the 11 dimensions, reduced in 1 dimension and increased in 7 dimensions.

A repeated measures ANOVA was performed with one within-subjects factor for the pre-intervention/post-intervention scores and another within-subjects factor for the 11 dimensions. This showed a significant effect of time ($F_{1,21}$ = 13.4, p = 0.001), but no time by dimension interaction ($F_{10,12}$ = 0.7, p = 0.7). This means that overall, there is a significant increase in the self-assessment scores between the pre-intervention questionnaire and the post-intervention questionnaire, with an average increase of 0.2 units over all the dimensions. There is no evidence from this ANOVA that this increase statistically significantly varies between dimensions.

In order to explore the results further, post-hoc tests were performed on the data from each of the dimensions. This was done using paired t-tests to compare the pre- and post-intervention results. The rationale for this choice of statistical analysis is described in sub-section 6.3.2. In this case data was only used for each dimension where the individual participant had completed both the pre-intervention questionnaire and the post-intervention questionnaire.

This meant that the comparisons involved 24 pairs of data for dimensions (A) and (B) and 23 pairs of data for the other 9 dimensions, which explains why there is a small amount of variation from the data shown in Table 6.1.

The results of these paired t-tests, which are illustrated in Table 6.2, demonstrate that in all but one of the dimensions there is an increase in score over time. (The negative number in dimension (B) 'Managing yourself' indicates a decrease in the score). However, this increase is only statistically significant for (E) 'Developing networks and building & maintaining relationships' (average increase is 0.5 units with a 95% confidence interval of 0.2-0.9) and (F) 'Working within teams and encouraging contribution' (average increase is 0.5 with a 95% confidence interval of 0.1-0.9). If a Bonferroni correction were to be applied to the numbers then neither of these increases would remain statistically significant, but as highlighted above there is a reasonable justification for not using the correction in this context of post-hoc tests.

Dimensions	Paired *t*-test	N
(A) Developing self awareness	$t_{22} = 0.6$, $p = 0.5$	23
(B) Managing yourself	$t_{22} = -0.4$, $p = 0.7$	23
(C) Continuing your personal development	$t_{21} = 0.5$, $p = 0.6$	22
(D) Acting with integrity and managing people	$t_{21} = 0.5$, $p = 0.6$	22
(E) Developing networks and building & maintaining relationships	$t_{21} = 3.0$, $p = 0.007$	22
(F) Working within teams and encouraging contribution	$t_{21} = 2.3$, $p = 0.029$	22
(G) Planning and applying knowledge & evidence	$t_{21} = 0.2$, $p = 0.8$	22
(H) Managing resources & performance	$t_{21} = 1.6$, $p = 0.13$	22
(I) Critically evaluating, ensuring patient safety and facilitating transformation	$t_{21} = 1.8$, $p = 0.088$	22
(J) Identifying the contexts for change and encouraging improvement & innovation	$t_{21} = 1.1$, $p = 0.3$	22
(K) Making decisions and evaluating their impact	$t_{21} = 0.6$, $p = 0.6$	22

Table 6.2 Comparison of pre-intervention and post-intervention self-assessment mean scores in the trainee committee group using paired t-tests

6.5 Conclusions

This chapter has described the quantitative comparison of the trainee committee cohort self-assessment questionnaire scores between the pre-intervention and post-intervention (12 months later) stages. 22 out of the original cohort of 32 (69%) participants fully completed the pre- and post-intervention questionnaires so the analysis covers these responses.

Comparison of before and after mean scores, using a global ANOVA, demonstrated a significant increase in the self-assessment scores between the pre-intervention questionnaire and the post-intervention questionnaire.

Additionally, paired t-tests demonstrated that in 10 out of 11 of the dimensions there was an increase in the mean score over time. However, this increase was only statistically significant for two of the dimensions:

(E) 'Developing networks and building & maintaining relationships'

(F) 'Working within teams and encouraging contribution'

It is important that any conclusion about the impact of the intervention is appropriately balanced. Although for 9 of the 11 dimensions there was no significant increase in the post-intervention scores there was a significant global increase, and also individual increases across 10 of the 11 dimensions. Overall, one has to conclude that there is a high probability that the first 12 months of the programme has led to some significant improvements in participants' leadership abilities.

6.6 Discussion

This discussion section is an opportunity to reflect back on, and scrutinise, the methodology used for the quantitative arm of the study. It also provides the chance to explore the comparative results and to review the conclusions that have been drawn from them.

Fundamental to this discussion is the choice of the methodology used in this study, and that is a theme that has been drawn out in different ways throughout the thesis. The systematic review of the medical leadership literature in chapter 2 illustrated a number of studies that had sought to evaluate the impact of leadership development interventions. All had in some way tried to measure the effectiveness of their intervention using pre- and post-intervention self-assessment surveys. The questions used were predominantly focused on Kirkpatrick Level 1 (reaction /

satisfaction) and Level 2 (self-reported learning / knowledge) evaluation methodologies (89) and in most studies there was no longer term follow-up conducted.

As detailed in sub-section 3.3.8.4, this study looked to use a comparator group within a 'quasi-experimental design'. The potential use of a control group, exposed to pre-test and post-test assessment was considered, but the pragmatic difficulties of whether a true control group would be able to be randomly selected, and then put through pre-test and post-test assessment, became a dominant factor. Although, as acknowledged in detail in chapter 3, this approach weakens the research design, provided due considerations are made to the recognised jeopardising factors, I believe that it is still an approach that can lead to some meaningful conclusions.

In sub-section 4.1.1 there is a detailed description of the thinking and exploration that was undertaken to attempt to find a more objective measure than solely a self-assessment survey used at the starting point and then end point of the study. Classic medical teaching suggests that a randomised controlled trial (RCT) would be the most effective way to understand the true impact of an intervention (204). However, in this situation the complexities and heterogeneity of the work-based intervention, the inability to 'blind' participants to which arm of the trial they were in, the inadequacy of a 'knowledge-based' assessment (as discussed in sub-section 4.1.1) and the logistics of managing a strict control arm mean that there was no prospect of this happening in this case. It is important to note that there are examples in the medical education literature where education-based interventions have been evaluated using a randomised controlled trial, and so it should come into consideration at the methodology planning stage of any educational research (205), even if ultimately other approaches are taken. As the medical education community develops a greater understanding around non-clinical learning this is undoubtedly an area that will benefit from further research. It seems likely that experimental designs with their origins in the social sciences (146)(149)(203) will continue to grow in their usage in this arena.

The debate around the appropriateness of self-assessment methodologies was explored in detail in chapter 4. As with the pre-intervention questionnaire it was hoped that the significant thought and time that went in to developing the content validity of the tool, the use of vignettes to help achieve context, the exploration of a relationship with the 'experiences log' and the extensive piloting would mean that the tool was as robust and effective as possible. Despite this, it was not a surprise that despite the global increases the results did not demonstrate statistically significant increases across all of the self-assessment scores. With relatively small

numbers of participants (only 22 fully completing both the pre- and post-intervention questionnaires) the study was under-powered. The constraints around the timings within this study meant that the post-intervention self-assessment was undertaken only 12 months after the intervention was introduced. With the intervention being much more focused on developing the leadership attitudes and behaviours that underpin many of the competencies within the MLCF (19), rather than just leadership-related knowledge, this was likely to be much too short a time frame within which to demonstrate widespread significant gains. The vast majority of the participants in the study have continued to be part of the trainee committee and so in theory it would be perfectly feasible to repeat the self-assessment questionnaire longitudinally at say 3 and 5 years post-intervention. The difficulty with interpreting these results would be the confounding effect of all of the different experiences (eg for some it would be starting as a consultant) that each individual may have encountered during that time period. This is an area where the weakness of the overarching study design, with no post-test assessment of the comparator / control group, is particularly problematic, although if a 5 year post-intervention assessment was made, this could potentially be rectified.

Despite the under powering of the study, the small sample size and the relatively short timescale for the leadership learning benefits to be realised, there were increases shown across 10 of the 11 dimensions, and a statistically significant increase when these were combined together using an ANOVA. Two of the dimensions did achieve statistically significant increases in their mean score in their own right; (E) 'Developing networks and building & maintaining relationships' and (F) 'Working within teams and encouraging contribution'. There is a clear justification, explained in sub-section 6.3.2.3, as to why the Bonferroni correction was not applied in this case. It is interesting that these two dimensions, which together represent all of the four elements from the 'Working with others' domain of the MLCF (19), were where the greatest gains were seen. Although the potential for learning across all of the domains was there, the majority of opportunities for experiential work-based learning from within the trainee committee were about working as a team, building relationships and developing networks – all hugely important competencies for doctors to be constantly looking to develop as they progress through their careers. Through the analysis of the experiences log described in sub-section 4.4.5 in chapter 4, the trainees who scored more highly on these cited experiences with patients, multi-source feedback, attendance at management meetings, participation in improvement projects, contribution to national working groups and patient safety groups and the teaching and supervision of colleagues as being important influences on developing leadership capabilities within these two areas. This is important as it gives some guidance to future programmes around the sorts of activities that

could be promoted and supported to give participants opportunities for meaningful leadership learning.

One other explanation for these increases in mean scores in these two dimensions could be that these are competencies that trainees will naturally develop over a year-long period though their normal clinical placements within the training programme. This would explain the 'one year on' increases in these areas. Without a post-test assessment of the comparator group, it is difficult to tell whether these increases might have been influenced by the history, testing, maturation and instrumentation jeopardizing factors described in chapter 3. However, reflecting back to sub-section 4.4.4, which explored the development of competencies as trainees progressed through training, there was no change in the dimension (F) 'Working within teams and encouraging contribution' over the different grades / years of training and no significant improvement over the dimension (E) 'Developing networks and building & maintaining relationships'. This suggests that the leadership development intervention described in this thesis may have had an impact of accelerating the learning in these two important areas.

As highlighted throughout the thesis, and in particular detail in sub-section 3.3.8, the key methodological design of this work was to take a mixed methods approach. Whilst the majority of the studies included within the systematic review relied only on pre- and post-intervention self-assessments there were two examples that sought to triangulate quantitative pre- and post-intervention self-assessment scores with qualitative data obtained through post-programme semi-structured interviews (91)(94). The methodological approach of running focus groups at the beginning of the leadership development intervention, and again after the 12-month 'end-point' sought to generate a deeper understanding and contextualisation of the quantitative results. The post-intervention qualitative data analysis is discussed in chapter 7.

Although the response rates to the self-assessment questionnaires compare very favourably to most other web-based surveys or questionnaires, almost one-third of the original participants (n=9; 31%) did not fully complete both the pre- and post-intervention self-assessment questionnaires. For two participants this was due to circumstances that led them to retire from the trainee committee mid-way through the 12-months. As with any consideration of the impact of non-respondents the concern is that their self-assessment answers, had they been captured, would have painted a different picture to the answers from the 69% of the group who did fully respond. Through the adoption of a mixed methods approach some of this concern is offset as all of the non-respondents participated in the pre-intervention focus groups (described

in chapter 5) and three of them also participated in the post-intervention focus groups (which are described in chapter 7). This meant that although there may be some responder bias within the quantitative data, this will have been significantly offset by the qualitative methodology, which gave the space for alternate viewpoints to develop.

Chapter 7 – Qualitative analysis of the end point

Chapter Overview

This chapter expands on the quantitative methodologies that have been used to measure the post-intervention (post-test) end point, by describing the focus groups that were run at the end of the first year of the programme. These focus groups were designed to capture a deeper understanding of the more qualitative aspects of the leadership learning gained from being a participant on the programme. The chapter also describes other themes and learning that emerged and includes a qualitative analysis of the use of the leadership multi-source feedback tool described in chapter 3.

7.1 Introduction

The previous chapter described the quantitative measurement of the 'end-point' (post-test) of the study, 12 months after the leadership development initiative had been introduced. This chapter takes a more qualitative approach by using focus groups to explore participants' perceptions and to build on the quantitative findings to assess the overall impact of the intervention.

7.2 Objectives

The main objective of this chapter relates back to the sixth overall objective of the thesis, namely to analyse qualitatively the effectiveness of the leadership development intervention experienced by the trainee committee cohort. There were also secondary objectives to qualitatively explore the use of the Leadership 360 (described in chapter 3) by the trainees who opted to use it, and to determine whether there were any aspects of leadership learning that were not well supported by the trainee committee.

7.3 Methods

This section details the qualitative methodologies used to assess the effectiveness of the leadership development intervention. The approach taken was for the most part very similar to the one used in the pre-intervention qualitative analysis. Where the methodology is identical, in order to avoid duplication the detailed illustrations of the background, context and rationale

of the methods outlined in chapter 5 have not been re-stated in this chapter. Where there are differences, their detail, and the rationale behind them, are described.

7.3.1 Data collection

Each of the focus groups was facilitated by the same pair of research assistants who had run the pre-intervention focus groups. As previously, a facilitation strategy was agreed with the lead researcher (RK) beforehand and an outline framework was developed. The research assistant (AR) moderated the discussions whilst the lead researcher acted as a non-speaking note taker (RK). As previously this role was to observe group dynamics and non-verbal interactions as well as noting down which statement was made by which individual. A digital audio recording was made for subsequent transcription. As before, and as discussed in chapter 1, participants were given an information sheet and gave their consent for their involvement and use of the data produced.

Focus group participants came from the trainee committee cohort, who were invited by e-mail to contribute on one of two different days. Three post-intervention focus groups were run. They had 6, 7 and 4 participants respectively (ie a total of 17) and lasted for between 35-50 minutes. This corresponded with the pre-intervention focus groups where there were 8, 7, and 6 participants respectively (ie a total of 21). Clinical duties for some trainees made it impossible to get all participants who wanted to contribute together at the same time so the first two focus groups were run on the same day (during a Trainee Committee update day), whilst the third was run just over 2 months later. The post-test self-assessment quantitative survey described in chapter 6 was sent out on the same day as the first two focus groups were run.

Each focus group was facilitated around the following four questions:

- Has working within the trainee committee influenced the development of your leadership and team-working skills? If so, how?
- Can you think about how this learning and development might influence your broader work within your clinical department in the hospital?
- Did you participate in the Leadership 360; if yes...did it help? If no....why not?
- Are there aspects of leadership learning that participating in the trainee committee has not supported?

The audio recordings were professionally transcribed into a digital typed document by an educational transcribing service, and as previously were independently checked before data analysis commenced. Any errors, in particular due to incorrect attribution of quotes, were corrected at this stage.

7.3.2 Data analysis

As with the pre-intervention focus groups, the methodology used to analyse the focus group data was 'framework analysis'. The same five key stages of familiarization, identifying a thematic framework, indexing, charting, mapping and interpretation were undertaken (see chapter 5 for detailed descriptions), with the aim of allowing themes to emerge, both from the research questions asked, and from the narratives of each of the participants.

The only differences in methodology were that, due to the lack of availability of researcher support at this stage of the study, these stages of familiarization were undertaken by the lead researcher (RK) on his own. This meant that the 'comparing of notes, categories and headings to reach consensus on the thematic framework' that had occurred in the pre-intervention framework analysis (see sub-section 5.3.4) could not happen in this post-intervention work. Although the same methodical step-wise approach was taken, the lack of discussion and debate could potentially lead to bias in the development of the thematic framework from this set of focus groups.

7.4 Results

This section describes the results and analysis of the focus group work. As outlined in the objectives section, and expanded on in the methods section above, the focus groups were designed to:

- Explore whether, and how, working within the trainee committee influenced the development of participants' leadership and team-working skills
- Consider whether this learning and development might influence the trainee committee participants' broader work within their clinical department in the hospital
- Explore the role of the Leadership 360 in supporting leadership learning
- Determine whether participants felt if there were any there aspects of leadership learning that participating in the trainee committee had not supported?
- Triangulate the quantitative data obtained from the post-intervention self-assessment survey

The 17 participants who were available to contribute to this part of the study were split into three focus groups as detailed in table 7.1. As with the pre-intervention focus groups, a count of the number of quotes from each participant that were adopted into the thematic framework was also made.

Group A Participant Code	*Group A Number of Quotes*	Group B Participant Code	*Group B Number of Quotes*	Group C Participant Code	*Group C Number of Quotes*
4	*5*	8	*1*	6	*4*
5	*3*	13	*1*	17	*3*
10	*6*	15	*7*	27	*10*
12	*6*	19	*3*	28	*7*
26	*6*	20	*10*		
30	*5*	23	*4*		
		25	*5*		

Table 7.1 Allocation of post-intervention focus group participants with illustration of the number of each participant's quotes used in the thematic analysis

Through the processes described in the methods section above three broad themes around leadership learning (usually called 'categories' within the framework nomenclature) arose from the framework analysis of the post-intervention focus groups. These suggested influence of the leadership intervention on:

1. Developing team-working, networking and collaborative skills and behaviours
2. Establishing new approaches to personal development and reflective practice
3. Understanding the importance of the wider context, change and the system within which healthcare is delivered

In addition, ideas emerged about the potential role for multi-source feedback tools in supporting leadership development, and finally, there were a collection of suggestions that focused on improvements that could be made to the design of this programme. As within chapter 5, quotes are attributed to individuals within each focus group and labelled with a 2-part code (eg GA; 12 or GB; 20) where the first two letters indicate which of the 3 focus groups they came from, and the number that follows being the individual's unique participant code.

7.4.1 Developing team-working, networking and collaborative skills and behaviours

The dominant theme that emerged from the post-intervention focus groups related to the experiences gained by almost all participants around team-working and the development of networks. This reinforces one of the main threads (199) of the quantitative findings described in chapter 6, where a statistically significant increase in self-assessed leadership capability was shown for the two dimensions 'Developing networks and building & maintaining relationships' and 'Working within teams and encouraging contribution'. These findings also fit in well with the theoretical frameworks on distributed (28) and shared (30)(31) leadership that were explored in chapter 1.

The trainees who participated in these focus groups highlighted how much they had learned from working with each other, and appreciating the value of each other's different approaches and leadership styles:

GA; 4	"It has been really interesting to meet a very dynamic group of people who have all approached a series of problems in a very, very different way; learning more about team working styles has had an impact on how I deal with everybody on all levels now"
GA; 30	"You learn from the other people in your group about different ways of going about doing things"
GC; 17	"I hadn't previously considered what I could learn from looking at other people"
GA; 12	"I think apart from the obvious networking, it is also knowing that you are not alone. Whatever problem I am having in my hospital, everyone else is having the same problem. I think what I have learnt is to listen to other people's ideas instead of banging my own head on the wall"
GB; 25	"Through talking to everyone and see what everyone is doing has really affected me and has made me a very different person; I realise now that I could change anything"

An important aspect of this learning came through the small-group exercises that explored the breadth of learning styles, personality types and approaches to team-working:

GC; 6	"Some people are quiet and effective...I still remember that from the first meeting"
GA; 5	"Talking about the Myers Briggs types is directly transferrable to the work place, as you can now try and work out what everyone else is and adapt to that"
GC; 27	"Respecting other people's way of leading, even if it is different from yours – I definitely got that out of the programme. I was very much of the opinion that if you were a quiet person it was because you did not have much worthwhile to say, but actually I have learned that this is absolutely not the case."

The trainees acknowledged that establishing effective team-working, especially with everyone committed in a number of different directions, was not necessarily straightforward, and required energy and focus to get the most out of it. For some it had been a surprise that this was such a key area of learning:

GA; 26 "the opportunity to work with people that you are not seeing all the time has highlighted quite a few challenges.....how we try to overcome these challenges have been perhaps part of our development"

GC; 28 "It has given me more of an appreciation of the give and take of team working"

GA; 26 "I think the opportunity to work alongside senior trainees or consultants in a non-clinical setting has been a really useful experience....and perhaps one that was not necessarily formalised at the start."

There was a strong sense of pride in everything that the trainee committee had achieved in its first year, and that had contributed towards a strong team focus that some participants felt would remain in the long-run. This was particularly eloquently explained in Group B:

GB; 15 "I think that the breadth of initiatives from all of us as a whole has been really impressive – the impact some of these have had as been amazing; the ability to recognise the need, initiate it and vault it into arenas far beyond what I had imagined...seeing that happen has given confidence and is empowering"

GB; 25 "I think empowering is the word for me; you just realise they have been waiting for people like us to step up to the mark and to just do it, and that is what has made so many things happen.

GB; 20 "I really like the idea that I have encountered a group of enthusiastic, dynamic, motivated people who I hope that through the rest of my professional life I will have a relationship with...and that makes me feel really inspired and excited as I think we can make a real difference in the years to come, even if this structure does not exist"

A number of the discussions illustrated a new found determination within the participants to act as catalysts for change. The focus was around empowering and motivating colleagues to believe that through working together to make small changes significant improvements could be achieved:

GB; 25 "I think that it has made me motivate everybody around me that small things can make a big difference; with lots of people putting things together you can change things"

GB; 19 "It has definitely motivated me to try to motivate my colleagues – before I would have been likely to have joined in with negative comments, but now I would be much more like to agree that it is rubbish but say 'lets all see if we can work together to improve it'..."

GC; 28 "Having seen how the trainee committee has developed, people are taking on projects and becoming leaders within their own small spheres of influence....and leadership roles."

GC; 17 "I think as you become more senior you have to stand back, be there, be supportive and work out how to give people that confidence and encourage that person so that you get the best out of them"

GC; 27 "It has made me better at getting my SHOs to really think about things, and to focus hard on giving positive feedback

Alongside this focus on motivating colleagues there was an acknowledgement of how valuable a learning experience being a participant on this programme had been, and that there was a moral responsibility and great value in ensuring that the leadership learning was taken back to the workplace to share with colleagues who had not had the same opportunity:

GA; 26 "We need to think about how a lot of these benefits and attributes that we are learning can also be passed onto other trainees."

GB; 13 "Also the importance of getting other people involved; Trust reps and roadshows were so important. It is not just about us having great ideas but taking them out to where we work, and bringing ideas back to here [the School] and I think that has happened quite nicely"

GA; 30 "It is taking back what we are learning to the people that we work with as well that is really important. We could just all be really selfish and go we have learnt all this stuff and just keep it all to ourselves but sharing it with others is good."

GC; 28 "I feel more confident about taking some of these ideas out to multi-disciplinary teams to try to make change happen"

Finally, the trainees highlighted specific areas of learning that were gained through their experiences on the trainee committee:

GC; 17 "Taking this onto the next step, I can see how this would help with managing conflicting ideas with a consultant colleague"

GC; 6 "I don't know how I would have managed to help facilitate people below me to learn stuff if I had not had a think about how to achieve this"

GB; 23 "I have definitely learned about being strategic about influence; who you speak to, how you approach individuals and how you find the right buttons to press"

In summary, participants described a wide range of learning relating to the development of team-working, networking and collaborative skills and behaviours. They acknowledged that this was both developmental for them as individuals, and would also influence the different teams within which they worked. There was also a widespread determination to spread their learning and experiences to a wider group of professional colleagues.

7.4.2 Establishing new approaches to personal development and reflective practice

In addition to the development of team-working, networking and collaborative approaches, it was clear that the experiences on the trainee committee led to significant personal development. One important area was the improved reflective practice of many of the participants:

GA; 12 "Dare I say it has broadened my horizons...has been a chance to sit and listen and learn something that is not clinical and I really enjoyed that"

GC; 27 "It has given me a clearer idea about my own leadership style, and has encouraged me to become more reflective"

GC; 6 "It is important to think about how we can use this [leadership] learning across other parts of our training and clinical practice. Doing the programme has made me more reflective"

Many of the trainees reflected on how their experiences over the preceding 12 months had allowed them to develop their own leadership styles and approaches.

GC; 27 "The course has also allowed me to learn how to not be a leader, but how to allow others to lead; rather than taking over....how to assist while still allowing them to lead"

GC; 28 "The trainee committee has allowed me to develop my own personal leadership style within the place in which I work"

GB; 20 "It has given me more appreciation of what there is to learn about being a consultant and manager and leader of a department; through the trainee committee it has made me appreciate that fact that when you become a consultant there are a whole load of other skills and knowledge that I need to take on board to be a good consultant. I have now taken on more of a personal responsibility to learn these things myself"

Another area that came up in all three focus groups was around the development of key personal qualities, such as self-awareness, a concept that, as described in chapter 1, is at the heart of the Medical Leadership Competency Framework (19) and many other leadership texts:

GC; 27 "I have started to consider much more about how I am viewed by other people; how I do I want to be seen by other colleagues and am I being seen in that way, and how can I act so that I am seen in the way I want to be seen?"

GC; 28 "I am less sensitive to changes and am more flexible about trying to do things differently"

GB; 19 "I agree that it has been really empowering. Now I am much more understanding that different issues are everybody's problem and I will do much more to improve things. As someone more junior I have realised that you can contribute, and you can make differences"

GA; 12 "I have definitely developed my listening skills; I have taken that back to the workplace and have developed that skill further"

In addition to the development of self-awareness, it is clear that enhancing trainees' confidence to take on projects and lead change was a positive outcome of the programme for some of the participants:

GA; 12 "I think that knowing all the stuff we are saying gave me more confidence"

GA; 10 "I took on, started and am still doing a piece of service development work at one of the hospitals I worked in - I think some of the skills we have learned, and taking the long view and really committing yourself to something that is beyond your clinical duties, has been really valuable"

GB; 20 "It has given me more confidence to know that I can make a real difference when I am a consultant, because I feel that the people who do the things that make a difference are the people who started off like we are now"

In summary, the focus groups highlighted three areas of significant personal development for a number of the participants. These related to becoming more reflective, developing more self-awareness and growing in confidence to take on projects outside direct clinical work.

7.4.3 Understanding the importance of the wider context, change and the system within which healthcare is delivered

The third area in which the focus groups illustrated significant learning arose through many of the participants developing a greater understanding of the wider system. This is another important leadership dimension covered by the Medical Leadership Competency Framework. It is also something that came up as a concern in the pre-test focus groups, meaning that it was something that could be added in to the leadership development itself. As acknowledged in chapter 3 this creates some significant methodological issues around contamination of the intervention (ie the focus groups in effect became part of the intervention), but it is positive to see from the post-test focus group that this addition was a positive one for the trainees. In addition there was an acknowledgement that change was more likely to be successful where the clinicians leading it had a broader understand of the context for change.

GA; 5 "Just understanding the broader context of where you are working so you can make sense of some of the things that actually happen to you at work and understand why they are happening"

GB; 20 "It has made me a lot more aware of what goes on behind the scenes, and how to influence, and who I need to know to make things happen"

GB; 15 "It has helped to broaden my horizons"

GA; 10 "I think it will equip us to be clinicians who take a wider view, and who as consultants will hit the ground running because we have been doing some of these things along the way. We will not feel so bogged down in day-to-day work but be able to see the bigger picture.... I think that is very valuable"

The trainees discussed ideas around change in all three focus groups. There was much enthusiasm to just get on and try to improve things, with the understanding that even small changes could be important. There was also the acknowledgement that this can be challenging and is a therefore an area that needs ongoing learning.

GA; 30 "I have got a much better understanding of how the system works, and the role of managers, and I have gone back to my department and changed things"

GB; 19 "It has made me more aware of the opportunities to make small changes – before I had not realised how important this could be"

GA; 10 "I think as an individual within the group there is still a lot for us to learn. I think navigating the bureaucracy of dealing with an institution such as the deanery, you become more aware of the actual challenges of effecting change. I am hoping I will just continue maturing in that sometimes it is easy to go in with lots of ideas but actually sometimes taking the long view is important"

Building on this theme of having a 'lot to learn', a number of the trainees also acknowledged how difficult change can be, citing limits, politics and hierarchies as significant barriers to effecting change:

GB; 23 "I have learned that in spite of all of that goodwill and enthusiasm, the NHS is a huge beast, and change is difficult ...and you have to work within certain limits, and that is a good lesson for many doctors"

GA; 4 "Learnt a lot more about the politics of the deaneries and the politics of the hierarchies of medicine which plays a major role in how we go about getting what we want done. ... with trying to effect and carry through and continue with change, it has been a real insight"

GB; 20 "With the example we have learned about Working Time Directive, even though we can't change it, there are ways in which we can use it to our advantage....and I think that is part of the learning experience; we have learned how to deal with it"

From the trainee's perspective, an important barrier to this broader learning arises from the frequent changes in posts that occur in most training programmes. This affects continuity and means that with trainees moving on every few months there is significant risk that trainees feel little ownership over work, projects are left incomplete and opportunities to learn from following things through are missed.

GA; 4 "I think continuity is something that we underestimate, especially on the junior level where you change Trusts almost every 6 months and you have no sense of the greater picture and the greater continuity. We have met lots of different people on the course; we came across medical directors, we came across NHS London people and got a picture of the greater strategic view. Having an awareness of what is outside the bubble and in what you do in a day to day basis and how you can direct your patients and help your patients - that has made a difference"

GC; 27 "Being on the trainee committee has given me lots of skills; what is difficult is putting them into practice when you are moving jobs so frequently. I can see so much I want to change, but more many things 6 months is not long enough"

GA; 26 "It has been a very useful lesson to know that if you are trying to set up something new in your hospital or in the community to actually engage people in it that are going to be involved in the longer term and give them some sort of ownership over problem solving and resolving whatever issues they have to resolve, is a much more effective way of getting things done"

In summary, this sub-section has described how the leadership development programme allowed a number of trainees to realise the importance of developing an understanding about change, and through doing so gaining a deeper understanding of the wider context of the system within which healthcare is delivered.

7.4.4 Multi-source feedback tools in supporting leadership development

The rationale, construction and results of the pilot Leadership 360 multi-source feedback tool, which was designed as part of this work, are described in detail in chapter 3. This subsection, which builds on sub-sections 3.3.7 and 3.4.5 details the focus groups discussions about participants' experiences of using the Leadership 360. The majority of those who used the tool found it very helpful in terms of generating feedback and helping them to further develop their self-awareness, although for some the feedback was not necessarily something they did not know already:

GC; 27 "I found it useful; I got a variety of people to do it...three people put down the same point that I had not considered before, which was something that I could work on. It was interesting to hear how the wider team perceived you"

GA; 10 "I got really useful feedback for that and I think it is on areas that as we have sort of said today that you don't normally get feedback on"

GA; 26 "The actual process of filling out the answers as well makes you think about leadership and was very useful as well"

GA; 30 "It would be useful to do it again because I have tried to change things about myself and what people wrote about me, so it would be useful to redo it and see"

GB; 20 "No-one said anything that I could use as point for development. I think when people sit down and write things, even though it is anonymous, they tend to just write nice things, and I can think of anything I read that made me think that is an area I am going to develop now...and that is why I didn't find it useful"

The discussions within all three focus groups identified that sitting down with a supervisor, mentor or senior colleague to go through all of the electronic feedback in the report was crucial

to the whole process. Where participants had not been able to do that the perceived value of the tool was much less:

GA; 26	"The actual feedback meeting with your supervisor ...and having some goals to achieve or ways of acting on whatever feedback you had is probably the most important bit"
GC; 28	"I liked the way that you could meet up with someone to go through the points....it gave me more focus. The feedback was informative"
GA; 30	"I think it is only a valuable tool, if you have got someone you can sit down with and talk about it and who will give you honest feedback. But if you are with someone who is just going to read through it who is just going to say 'oh haven't you done well' or whatever, is not helpful at all. Maybe there needs to be more structure about who is going to give feedback?"
GB; 15	"Without talking through the feedback with anyone....taking home suggested changes from the Leadership 360 was difficult and challenging"
GB; 25	"One person said "sometimes she can be abrupt".....or something like that, and that is all I can remember from my Leadership 360. I realise that if I had sat down and discussed it ...then something much more useful would have come out, and then there would have been something to work on"
GB; 15	"The stumbling block was finding the time and the right role-model to go through it"
GB; 23	"To draw out positive themes with any kind of clarity or distance is difficult to do when it is really personal, and so I didn't find it that helpful because I kind of left it alone and did not go through it with anyone"

One unexpected source of feedback on the value and quality of this Leadership 360 tool, that arose during two of the focus groups, came from colleagues of the participants who had been asked to contribute. The breadth of areas covered by the tool seemed to be the key theme that provoked comment:

GA; 10	"I did get some feedback from the people who filled it in that was very positive as well. One of the PICU consultants said it was very good survey. They have done a lot of these things and they were quite surprised and found it particularly useful because it dealt with issues that you do not normally get to talk about"
GB; 15	"The feedback I got was that the people who filled it in said that they really enjoyed it because they felt that it was a challenge to think and reflect in a way that we don't normally do; they as assessors found it more useful than the normal 360"

In addition to the importance of ensuring that participants go through the feedback with a colleague afterwards, a number of other comments and suggestions were made that might improve future iterations of this or similar leadership multi-source feedback tools. Some of these related to the timing of when the tool was available, others with technical aspects of the tool, and others to the difficulties in getting enough people to give feedback:

GB; 20 "Not being able to print it made it difficult"

GC; 27 "I think it was not all that well anonymised; you could see when someone had responded even though you could not see their name and that might mean you could work out who they were"

GB; 8 "It was quite a laborious questionnaire"

GB; 23 "It was too close to the other assessments which meant I had to persuade people to complete it for me"

GA; 4 "I think timing is the big issue as a lot of these things tend to happen around changeover time which is the worst possible time to get people to assess you - yes you can get hold of people from previous jobs to do so but this is not easy"

GC; 17 "I was somewhere where it was difficult to get anyone to do anything at all, even the compulsory stuff, so to do anything extra was too difficult"

GC; 6 "I was an ST2 in a tertiary neonatal unit, bottom of the pile, so didn't really want to send around a 'do you think I am a good leader?' questionnaire"

In summary, participants used the focus groups to give considerable consideration to the value of the Leadership 360 multi-source feedback tool. For everyone who managed to participate, the crucial determinate as to the value of the experience was whether or not they were able to go through the feedback with a supervisor, mentor or senior colleague. Where this had occurred the tool provided helpful developmental feedback. This crucial point about feedback fits in with much of what is known about multi-source feedback tools in the healthcare literature (151)(154). The need for trainees to complete other compulsory multi-source feedback assessments at a similar time was the key barrier to using the Leadership 360 for a number of participants.

7.4.5 Suggested improvements to the design of this programme

Throughout the three focus groups a number of discussions generated ideas that could be used to improve the design and running of the programme. One important dimension that arose from the Group A focus group was the recognition that participating in this leadership development initiative had been very valuable, and that there was a responsibility to share ideas and resources with those who had not been able to participate:

GA; 4 "I feel that we have been privileged to get what we have had out of it and kind of feel for all my colleagues that did not jump on the bandwagon and be in it as well"

GA; 10 "I think maybe there is room for doing a bit more of sharing of resources…..maybe online or something"

GA; 12 "At the end of a session to have a few minutes to write some reflective notes….just so that the people who have missed out will know what they have missed out on and will be able to go onto that website and have a look at it"

The opportunity to receive some face-to-face teaching on leadership-related topics during the programme was well received, and a number of other suggestions for specific sessions were put forward:

GB; 25 "Having a couple of hours of leadership teaching, where we had a speaker and discussions, was totally brilliant"

GA; 12 "I do not think we have actually touched on how to deal with a like a difficult situation, like conflict resolution, apart from learning it in our teams.... something formal would be good"

GA; 5 "When you are back in your own Trust and you're trying to facilitate a session or lead some teaching or something, some of those skills might be really, really useful....so maybe some teaching on teaching styles and how to encourage participation would be really useful"

GA; 10 "It would be nice to have some support about managing your juniors.... and knowing yourself; they are common skills but maybe doing it in a sort of transparent way with clinical scenarios"

Another useful learning activity that was suggested was that of 'shadowing' or finding the opportunity to sit in on the School of Paediatrics executive committee:

GA; 10 "One thing I would like to do sometime is to be able to sit in on the executive committee and see what happens there.... I know there is a probably representation from us which is fine but I think as a learning opportunity that would be really useful"

One of the problems highlighted was the difficulty of being able to attend all of the meetings. For some this led to a feeling of anxiety that they were not contributing a much as they would have liked, whilst others explained that they were concerned that participating in trainee committee would not be well perceived by colleagues back at the hospital. Others acknowledged that they were not able to attend everything but used the strength of the communication within the sub-groups, and the trainee committee network, to offset this:

GB; 20 "I found it hard to find the time to come to as many meetings as I would have liked, and I know that I am not contributing as much as I would like. I suppose there will be different times for everyone where we are able to contribute more or less"

GC; 27 "I could only go to the meetings sporadically, so I don't know what I missed"

GC; 6 "I erred on the side of caution because I did not want to be perceived as someone who was 'swanning off' to things that were not thought to be that important"

GB; 15 "In the way in which it runs, isn't that almost accounted for? We do operate as a network who know each other....and that is the power of it that means it can be much more flexible in these manageable sub-groups"

The trainees highlighted the role of consultant support and leadership within the sub-groups, and explained some specific aspects of this that they perceived to be particularly important. The experience of the trainees around this had been variable:

GB; 15	"The only disappointing thing has been the consultant that has led one of the groups….in terms of knowing current topics and agendas – we needed to be pointed in the right direction with some of these agendas and that has been sorely lacking in our group"
GB; 20	"He has been amazingly supportive and proactive; he has been amazingly involved and has not left all the work to us. I think a lot of what we have achieved is down to his leadership"
GB; 20	"He is very good at allocating roles and setting deadlines, and being realistic about what we can achieve. He is also very good at including us, he asks what we think. He is very much a model of a consultant who wants to have trainees involved"

As a footnote, the trainees in the third focus group described how they had found participating the focus groups a helpful way to pick up new ideas and reflect on where they needed to take their learning next. This evidences the thinking that these focus groups need to be thought of as part of the intervention:

GC; 28	"I have found these focus-groups useful in terms of getting ideas off other people. I remember the first focus group, and the discussions about what makes a leader certainly changed my views on that"
GC; 27	"I think that what is also important is to have ongoing training, even if it is similar to what you have had before..... today has been a reminder to go and put some of the skills I have learned back into practice"

In summary, participants described a number of ideas that could be fed back into the design of a future iteration of this programme. These included considerations as to how they could better share learning resources within the trainee committee and with colleagues outside it, other ideas for face-to-face teaching topics, and the role of shadowing in supporting leadership learning. The trainees also discussed the difficulties in attending all of the sessions, and the importance of consultant leadership and enabling of the subgroups. Finally there was an acknowledgement by one group that taking part in these focus groups had been a valuable way to further progress their learning.

7.5 Conclusions

It is clear that overall the trainees found their participation in this leadership development initiative to have been extremely rewarding, both in terms of their personal learning and development, but also with regards to everything that they achieved as a trainee committee. The trainees' main areas of learning were around establishing new approaches to personal

development, in team-working, networking and encouraging the contribution of others, and in broadening their understanding of the wider healthcare system.

Those that had managed to complete the Leadership 360 and have a face-to-face discussion about it with a supervisor or mentor, described how much they gained from it. In addition, the trainees also had a number of ideas around how future iterations of the programme could be improved.

7.6 Discussion

The methodology used in this chapter followed that of chapter 5, and was underpinned by the methodical steps of 'framework analysis'. It is important to acknowledge that in this post-intervention work, although the same steps were taken, only the lead researcher was involved. This meant that the advantages of discussion and challenge between 3 different researchers, that had added richness to the pre-intervention work, were absent here. There is a risk that this could introduce bias into the analysis. A second concern could relate to the reducing number of participants. Although one year on the numbers of participants in the three focus groups had fallen from 21 to 17, which might lead to an increasing impact of non-participation bias, they covered a wide range of ideas and opinions that met all of the objectives of this chapter. It was clear that many of the participants felt a good deal of pride in everything that they had achieved, in what was a relatively short period of time. In addition, as a result of their experiences on the trainee committee, many of them felt very energised and ready to drive forward improvements and change in their place of work.

Importantly, the focus group analysis follows on from the threads within the quantitative results from chapter 6 which demonstrated that in 10 out of 11 of the dimensions there was an increase in the mean score over time, with statistically significant increases for two of the dimensions, (E) 'Developing networks and building & maintaining relationships' and (F) 'Working within teams and encouraging contribution'. Many trainees described how much they had learned from each other and the dominant theme that emerged related to the experiences gained by the majority of participants around team-working and the development of networks. This is an important finding and although impossible to put a definitive 'weighting' onto it, strengthens the opinion that the increases in the self-assessment scores for dimensions (E) and (F) demonstrated in chapter 6 represent genuine improvements in leadership abilities of the participants. The qualitative analysis also described what appeared to be a fundamental shift in some participants in terms of an improved understanding of the perspectives of others. These are important areas that also relate back to the theories around

transformational leadership, and shared leadership, discussed in chapter 1, as well as a number of the team-based leadership development programmes, such as those highlighted in the Evans (85) and Stoller (86) papers described in chapter 2.

Many of the trainees talked about how this learning would be invaluable to them as they moved forward with their careers in the years ahead. As well as developing new, more reflective approaches to their own personal development, for many trainees this programme was an opportunity to learn about the wider healthcare system. As with the 'end-point' quantitative survey described in chapter 6, this focus-group work, which took place 12 months after the programme was initiated, will by definition have only asked about relatively early changes in knowledge, understanding and perspective. For any sort of formal analysis of longer term impact on individuals and groups of trainees in terms of behaviours, attitudes and achievements, the cohort would have to be revisited with further measurements a number of years down the line.

With the opportunity to undertake the Leadership 360 only being taken up by a minority of the trainee committee, it is difficult to make too many conclusions from its design and use. It is important to remember that it was entirely optional for the reasons discussed in chapter 3. However, this qualitative analysis highlighted that those who had used it as it had been designed gained a significant amount of learning and insight from the experience. The spontaneous feedback from consultant assessors as to how impressed they were with the depth of questions covered by the Leadership 360 was a very welcome benchmarking against other, predominantly commercial, multi-source feedback tools that they had come across. They also commented on how the vignettes were helpful in contextualising the question being asked. The trainees were clear in their recommendation that participating in this type of multi-source feedback was only worthwhile if there was a focused debrief of the results with an experienced supervisor or senior colleague, and this is important learning that can be adopted by other leadership development programmes using these methodologies. At a time when the medical profession in the UK is developing its processes of 'revalidation', the General Medical Council's mechanism for regulating licensed doctors, this also offers an important lesson in terms of getting real value out of appraisal. The experience of the trainees with the Leadership 360 also highlighted a key issue where there are a number of different assessments and feedback tools, some of which are compulsory components of trainee schemes, competing for the time of trainees and consultant assessors. When it came to it, even for trainees who could see great value in participating in a leadership-focused multi-source feedback tool, the mandatory Royal College assessments took precedence. This learning could have significance

for the widespread use of the multi-source feedback tools made available to NHS workers alongside the launch of the NHS Leadership Framework in 2011 (158).

The trainees also made a number of comments and suggestions which could be used to improve future iterations of this programme. Throughout the reading of the wider leadership literature, the systematic review of the medical leadership literature and the self-assessment results the value of using real work-based opportunities to drive learning has been an important thread. However, as discussed in two of the focus groups, a number of the trainees were worried about not being able to attend some of the trainee committee meetings and events. Even in designing a work-based programme of learning, where it was acknowledged from the outset that not everyone would be able to attend everything, this caused anxiety amongst participants. The suggestion of using IT solutions such as a web-based forum or social media to support communication between the group may help to offset these concerns. The discussions around the importance of consultant leadership in enabling and supporting the trainees as they set their vision for their subgroups aligned with the value of role-modelling ascertained from the quantitative work in chapters 4 and 6. Finally, there was also a brief discussion about the value of shadowing senior colleagues and having the opportunity to sit in on executive-level meetings (in this case the School of Paediatrics executive). This idea, which is expanded on in the 'next steps' sub-sections of chapter 8, went on to strongly influence my work in the development of the shadowing component of the 'Paired Learning' leadership development initiative (206)(207)(120).

Chapter 8 – Summary discussions

Chapter Overview

This chapter focuses on summarising and analysing the findings from the quasi-experimental mixed method approach to evaluation described in the previous six chapters, and then contextualising them within the wider leadership literature from chapter 1 and the medical leadership systematic review in chapter 2. It examines the strengths and weaknesses of the research, and synthesises the findings whilst exploring the wider implications of the work. The chapter contains some reflections on my personal learning from this project, before moving on to consider possibilities for longer term follow-up and the applicability of the findings to other groups. The chapter concludes by summarising the key conclusions and discussing how the learning from this project might be implemented within other leadership development initiatives.

8.1 Summary of findings

In order to synthesise the work that is described in each of the preceding chapters it is worth reflecting on each of the objectives of the project that were set in the planning phase. These objectives were to:

a) Perform a focused systematic review of the medical leadership literature, and to analyse key learning and themes around medical leadership development initiatives.
b) Design and implement a leadership development initiative to support the learning and development of a newly established postgraduate trainee committee, and design a mixed methods research strategy to evaluate it.
c) Develop a self-assessment tool to measure the leadership capabilities of the trainee committee cohort prior to the implementation of the leadership development intervention, and to compare these results with a comparator group of trainees.
d) Explore participants' ideas around leadership to gain a deeper, qualitative understanding of the baseline leadership learning of the trainee committee cohort, prior to the implementation of the leadership development intervention.
e) Evaluate quantitatively the effectiveness of the leadership development intervention experienced by the trainee committee cohort.
f) Analyse qualitatively the effectiveness of the leadership development intervention experienced by the trainee committee cohort.
g) Summarise the findings from the project, highlighting areas of personal learning, potential wider implications of the project and suggestions for future work.

The systematic review of the medical leadership literature described in chapter 2 confirmed that at the time of the review there had not been any other studies of evaluated work-based leadership development programmes for doctors previously described in the literature. There were papers that provided ideas about work-based learning in other settings, and these were used as empirical evidence to help shape the intervention. The review highlighted the fact that amongst the huge number of leadership related papers that have been published only a very small number have included a well-defined intervention and some sort of evaluation of the impact on the learners. Of the published work with well-defined interventions, the majority were evaluated by pre- and post-intervention self-assessment surveys alone; the use of qualitative methodologies and long term follow-up was only seen in a handful of studies. The included studies emphasised the value of linking leadership learning to real work and the importance of programmes having the input of senior leaders. They also tended to focus on supporting the development of self-awareness and team-working competencies. This learning helped to inform the evaluation strategy of this study; the inadequacies of most studies to be able to draw any meaningful conclusions from their stand-alone self-assessment metrics re-iterated the potential value of taking a mixed methods approach to the evaluation in this project.

The focus of the leadership development initiative was to support the learning and development of the newly established postgraduate trainee committee, and chapter 3 details the design, planning and implementation of this initiative. The application process to the trainee committee led to the formation of strong cohort of 32 trainees. Ethical approval was successfully gained for the implementation of the proposed leadership development initiative and underpinning evaluation strategy and all participants were given information sheets and consented to participate in the key areas of the research (eg the focus groups). A leadership self-assessment tool was developed to quantitatively measure the pre-intervention leadership capabilities of the trainee committee cohort. This tool gained content validity and an understanding of its feasibility from its relationship to the Medical Leadership Competency Framework (MLCF) (19), from the embedding of case vignettes with each question and through detailed piloting. Within the tool, an interval 5-point semantic differential scale was designed, based on published scales of ability and expertise (183), and then piloted. On reflection there were some methodological weaknesses with the design of this scale, which are explored in more detail in sub-section 4.3.1.3 and in the extended discussion section 4.6. The validity of this self-assessment tool was somewhat strengthened by the demonstration of a correlation between the scoring and the number of experiences that the participant had encountered to develop those capabilities. Analysis of the data also showed that certain

specific previous experiences were more likely to influence the development of self-assessed leadership abilities than others.

Prior to the commencement of the programme this self-assessment tool was completed by the trainee committee cohort and almost 300 other paediatric trainees, who were acting as the comparator group. Statistical analysis showed that at this stage (prior to the introduction of the leadership intervention) there was no significant difference in leadership ability between the cohort and comparator groups. Importantly, more detailed analysis of the data demonstrated that the more senior trainees only self-assessed themselves significantly higher than more junior colleagues in one leadership dimension (B) 'Managing yourself'. This suggests that within the standard training programme it should not be expected that the leadership capabilities of trainees will necessarily develop in time as trainees move up through the grades. With an inherent methodological weakness of the overall project in having no post-test of the comparator group (as discussed in sub-section 3.3.8.4 and section 3.6), this helped support the suggestion that any post-intervention increases in the trainee committee cohort identified with this tool could potentially be attributed to the impact of the programme.

Chapter 5 described the framework analysis of three focus groups to explore the trainee committee participants' ideas around leadership. This provided a deeper, qualitative understanding of the baseline leadership learning of the trainee committee cohort, prior to the implementation of the leadership development intervention. This allowed for integration with the baseline (pre-test) quantitative analysis outlined in chapter 4. Four themes emerged: the importance of leadership qualities and leadership behaviours in doctors as leaders, ideas around opportunities that would help the participants develop as leaders and an understanding of the value of broader aspects of leadership training and learning. These themes were used as a further input to shape the content and learning outcomes of the leadership development intervention.

After 12 months of the leadership development intervention 'end point' (post-test), or post-intervention, measurements of the trainee committee were made. The quantitative evaluation, which used an adapted version of the pre-intervention leadership self-assessment, allowed for a comparison of before and after mean scores of the trainee committee cohort using a global ANOVA test. This demonstrated a significant increase in the self-assessment scores between the pre-intervention questionnaire and the post-intervention questionnaire. In addition, paired t-tests demonstrated that in 10 out of 11 of the dimensions there was an increase in the mean score over time. However, this increase was only statistically significant (pre-Bonferroni

correction) for two of the dimensions: (E) 'Developing networks and building & maintaining relationships' and (F) 'Working within teams and encouraging contribution'. Overall, this leads to the conclusion that there is a high probability that the first 12 months of the programme led to some significant improvements in participants' leadership abilities.

The impact of the intervention was also assessed with the focus group-based qualitative analysis described in chapter 7. The trainees who participated in these focus groups were clear that their participation in this leadership development initiative had been extremely rewarding, both in terms of their personal learning and development, but also with regards to everything that they and their colleagues had achieved as a trainee committee. They described their main areas of learning as being around establishing new approaches to personal development, team-working, networking and encouraging the contribution of others, and through broadening their understanding of the wider healthcare system. In addition the trainees who had managed to complete the Leadership 360, and have a face-to-face discussion about it with a supervisor or mentor, described how much they gained from it. The trainees also had a number of ideas and reflections about ways in which the programme could be improved – these included comments about the sharing of learning both within the cohort (for those who could not attend specific sessions or meetings) and with colleagues who were not formally part of the trainee committee. They emphasised the need for consultant leadership, support and enabling of work-based learning, and also suggested the potential to learn from shadowing or sitting on activities and meetings that they would not normally be exposed to.

From these findings, and despite some of the methodological weaknesses of the study, it can be concluded that clinical leadership skills and capabilities *can* be learned by trainee paediatricians. In addition to the clearly demonstrated impact on many of the individuals involved, and the personal learning and development for me as an educational researcher, this thesis has generated new evidence that can be used to influence and drive the design of other leadership development initiatives in the future. The remainder of this chapter is devoted to exploring this in more detail.

8.2 Analysis of findings

Throughout the thesis, in the methodology and discussion sections of each chapter, there has been a detailed on-going analysis of the strengths and weaknesses of the research. This has included the rationale behind each of the chosen methods, detail as to why others were rejected and an open narrative around how different problems and inherent weaknesses could be mitigated against or needed to be acknowledged. This has been contextualised in the

theoretical frameworks of the leadership literature review described in chapter 1 and in the empirical evidence gained from the medical leadership systematic review detailed in chapter 2. The key strengths and weaknesses are summarised below:

8.2.1 Strengths of the research

The section above describes some valuable gains for the trainees who participated in the programme. The hope was that the leadership development intervention, based around the work of the trainee committee, would give the participants the opportunity to experience and reflect on followership, but also to recognise their important role as leaders within the healthcare system, despite still being labelled as 'junior' doctors. The findings suggested specific influence of the leadership intervention on three key areas:

1. Developing team-working, networking and collaborative skills and behaviours
2. Establishing new approaches to personal development and reflective practice
3. Understanding the importance of the wider context, change and the system within which healthcare is delivered

In addition, ideas emerged from the trainees for the potential role of multi-source feedback tools in supporting leadership development. The programme was not designed to be modular, but instead was built around the idea that being an active participant in the trainee committee, and contributing to specific project work, would together lead to leadership learning. The workshops and opportunity to participate in the Leadership 360 were put in place to support this. Perhaps unsurprisingly, trainees were not particularly specific about which aspects of the programme contributed to particular areas of learning. Therefore, as with all of the multi-component interventions considered in the medical leadership systematic review in chapter 2, it would be difficult to value the impact of one component of the intervention over another. One important reflection is that the opportunity to work and learn as a team, be it the committee work, the sub-group project work, or indeed the involvement in the focus groups, was highlighted by many trainees as being important.

The research was underpinned by a grounding in key leadership theories (outlined in chapter 1), the use of the Medical Leadership Competency Framework (MLCF) as a guiding structure, and a comprehensive systematic review of the medical leadership literature (outlined in chapter 2). The principles that the two BEME papers (73)(78) offered were extremely helpful and enabled the development of a methodological framework that could be closely followed. The efforts to review lists of references, to explore studies that cited the retrieved papers, to

revisit the searches with an August 2012 update and to explore the grey literature all added to the strength of this work. This gave important context to the research and provided ideas for the development of the intervention and the evaluation methodologies. It also confirmed that there were no other healthcare-related work-based leadership development initiatives, underpinned by a comprehensive evaluation strategy, in the published literature. This meant that both the intervention and the evaluation methodology (with the development of the self-assessment leadership ability scale) were novel and of potential interest to other educators, clinicians and researchers working in this area.

The decision to proceed with a work-based intervention was an important aspect of the study, in terms of the practicalities of the programme, applying learning and developing an understanding of the wider system. This also played into supporting participants in having the opportunity to gain many of the competencies described in the Medical Leadership Competency Framework (1). Whilst there may be value in classroom-based leadership programmes clinicians of all levels find it increasingly difficult to free-up blocks of time from their clinical duties and so attendance can be difficult. The flexibility of this programme and the ability to dip in to the different opportunities at different times meant that this was less of a problem. In addition, the ability to apply the leadership learning to real situations within the work-place gave greater context and a deeper understanding and perspective. The post-intervention focus groups alluded to the achievements of the trainee committee – a work-based leadership learning approach that also has specific outcomes related to better efficiency, patient safety, improved quality of care or patient experience is much more likely to gain executive support within healthcare organisations for implementation and spread.

Although not without compromises and practical difficulties the mixed methods approach to the evaluation was a significant strength of the study. As detailed in the systematic review in chapter 2 there is a huge amount published around leadership development, but relatively little of it has been evaluated in any great detail. Chapter 4 describes the different steps that were undertaken, through the content validity gained from the MLCF and other sources, the use of case vignettes, the piloting, the development of the semantic differential ability scale and its correlation with previous experiences, to make the quantitative assessment as strong as possible. The use of a comparator group to establish the baseline 'starting point' and the generally high response rates across all of the different 'experiments' were also helpful. Perhaps most important of all, and unique amongst the previously published research in this area, was the use of the two sets of focus groups as a qualitative mechanism to make sense of the self-assessment results. Chapters 5 and 7 describe in detail the methodology used in the analysis of the focus group data, and the subsequent results. Although sometimes criticised as

being somewhat cumbersome, the method used, framework analysis, gave both depth and structure to the analysis. In addition to providing evidence that aligned with the quantitative results, the different analyses from the focus groups generated ideas and reflections that could be used to shape future research.

Although there were no particularly contentious ethical dilemmas within this study, the decision to go through the process of application for ethical approval ensured that all potential ethical issues were considered and also provided the study with some early peer-review.

8.2.2 Weaknesses of the research

As described above, an overall aim of the study was to develop an initiative that was highly flexible and pragmatic in acknowledging that there would never be occasions when all 32 individuals on the programme could come together in formal leadership learning sessions. The intervention was also very low-cost and therefore, if shown to be of value, would have a good chance of long-term sustainability and spread. The main difficulty of this approach was from a research perspective, where the 'blurring around the edges' of which experiences were inside the defined leadership development intervention and which were not, was complex. The problems with this were four-fold. Firstly, there was a risk that any improvements or change following the leadership development intervention may not be clearly picked out by the research methodologies. Secondly, any changes picked up by the evaluation may actually be attributable to some other factor outside the influence of the leadership development intervention. Thirdly, aspects of the evaluation methodology (eg the pre-test focus groups) may have had an impact on the results because of interaction and sequencing issues, and so in essence it can be argued that they became part of the intervention. Fourthly, when thinking about sustainability and spread, reproducing exactly the same experience for other cohorts would be extremely difficult. Managing these concerns, whilst ensuring that the leadership learning available to the trainees was highly flexible and iterative, was a major challenge for this study.

Although the systematic review described in chapter 2 was comprehensive, and was designed to take a broader approach than most medical systematic reviews, as with any search of this nature, the difficulty in achieving full coverage has to be acknowledged. Not all databases or possible search terms were covered and so there may have been key studies that were not found. The efforts to explore the grey literature, described in sub-section 2.3.4, are likely to have only found a small proportion of the information available. It must also be acknowledged that the decision to stick to leadership development interventions in doctors and medical

students may have restricted this search for empirical evidence. It was therefore important to ensure that all possible learning was gained from the papers and information that were found in the review, and so efforts were made to incorporate any relevant findings or learning into each of the other chapters.

As described in detail in chapter 3, and also mentioned in chapters 6 and 7, the use of a comparator group within a quasi-experimental mixed methods evaluation design is not without controversy. Although the use of a control group, who would also be subjected to both a pre-test and post-test assessment, was considered in this project, the pragmatics of selecting a true control group, and then putting them through pre-test and post-test assessments proved to be impossible. It has to be acknowledged that this important practicality weakens the research design, and therefore limits the conclusions that can be made from the findings. It means that the potential impact of a number of different jeopardizing factors (146), including history, maturation, testing, and the effect of autonomous changes in the measuring instruments between the pre-intervention test and the post-intervention test need to be considered. Although the attempt to use a comparator group of trainees from multiple year groups, as a way of offsetting some of these problems, had partial impact in mitigating some of these factors, this remains an inherent weakness of the project.

As discussed in sub-section 3.3.8 this approach also creates interaction effects that lead to potential issues with external validity. These are interactions that occur between the intervention and the selection process used to select the cohort group, and between the intervention and the testing processes. These factors can lead to difficulties with generalisation of the results to other settings, which could have an important effect on how conclusions from this study are interpreted.

The choice of using a self-assessment tool to collect data for the pre- and post-intervention quantitative analysis is another controversial area. The key concerns and reasons for the decision-making around this part of the study are detailed in chapter 4. Sub-section 4.1.2 highlights that although there are strong proponents of their use within leadership development there is good evidence of the limitations of self-assessment when used to assess the abilities or performance of medical professionals(164)(165)(166). The suggestion is that doctors are likely to over-estimate their self-assessment scores when compared to existing objective, externally generated markers of performance. There is also evidence to suggest that the least competent are also the least able to self-assess accurately, often significantly over-estimating their own level of performance (167). Other evidence however, identified that the accuracy of self-

assessment tools can be enhanced by feedback, and also by providing explicit assessment criteria and guidance on what level of performance should score what mark (168). The comprehensive approaches to mitigate against this, through the careful design and piloting of the self-assessment tool, which are explained in detail in chapter 4, were therefore very important. The discussion in chapter 4 also highlighted a methodological issue with the semantic differential ability scale that was developed for the self-assessment tool. It is possible that because the extremes in the 5-point scale were not strict opposites, there was no defined neutral mid-point and there was more than one concept within the scale may have skewed different participant's perceptions of the questions being asked.

Another potential weakness of the study relates to the relatively small number of participants, although compared to many of the studies in the published literature a study cohort of 32 is reasonably large. With the quantitative analyses this did mean that the study was relatively under-powered. 'Drop-off' in terms of people starting to answer the questionnaires and then not completing them was not a big issue within the cohort group, although it did reduce the comparator group from 336 to 289 participants. A significant effort was made through testing of functionality, individual messaging, reminders and ease of participation to keep response rates and completion rates as high as possible. Overall this approach was successful with the response rates to the two self-assessment surveys and both sets of focus groups, each of which are detailed in chapters 4, 5, 6 and 7, remaining high when compared to equivalent studies. There remains, however, the concern that the non-respondents to the survey and non-participants in the focus groups would have given different responses to the participants who did contribute. Response rates and the issue around non-respondents were perhaps most significant an issue with the post-intervention 'end-point' focus groups, where effectively only 17 out of the original cohort of 32 trainees were able to participate. Although many of the non-participants had continued to contribute very fully to the trainee committee, and had sent apologies for the focus groups due to clinical commitments, it would have been interesting and important to have explored their views and perspectives too. Due to logistics and constraints of time this study did not seek to investigate this further, but this might be an important, if difficult, area to explore with future research.

A further related problem was that of 'self-selection'. Although the trainee committee cohort were shown to have no significant differences in their pre-intervention self-assessment scores to the group of almost 300 paediatric trainees who acted as the 'comparator' group, there is a weakness to this argument. By nature of the fact that they were the individuals who responded to and completed the leadership capability survey, this 'comparator' group, although coming

from the same paediatric trainee population as the trainee committee cohort, were in some ways 'self-selected' and were not strictly randomised to be comparators or indeed controls.

Despite the strengths of the structured, 'framework' approach (193) to the analysis of the focus groups there are some inherent weaknesses in this methodology and how it was used in this research. Although the analysis of the data was both step-wise and rigorous it is still susceptible to inherent biases of interpretation and emphasis that are held by the researchers. Although the pre-intervention focus groups were able to use a combination of three researchers to help protect against these issues, the post-intervention focus groups did not have the same breadth of independent analysis followed by discussion, and so may have been more subjected to these biases. The fact that the lead researcher (RK) was intrinsically involved in the development, co-ordination and running of all aspects of this research, and in particular the evaluation, could also mean that intrinsic biases could affect the way participants answered and the way in which the data was analysed. The only way to completely avoid this would for the programme of evaluation to have been independently run by a team of external researchers, and as with the majority of health-related development programmes this was not logistically possible.

Finally, with the nature of the work-based learning and development that this programme hoped to achieve, it was perhaps unrealistic to expect wholesale significant development over a 12 month period. Almost 90% of the trainee committee remained in their roles beyond 12 months, and in the context of a paediatric training programme that runs for at least 8 years, this type of learning seems most likely to need to have the space and opportunity to develop over a longer period of time. The constraints around the timings of this thesis were the factors that set this time scale, and if it were possible it would be of great interest to go back to the trainee committee cohort to look at the longer-term impact of the intervention. As stated earlier in the thesis, despite the inherent logistical difficulties, this might also open up the opportunity to take a different methodological approach, by undertaking post-test self-assessment surveys in both the trainee committee cohort group and the comparator group, which would then in effect become a control.

8.3 Next steps

This section explores some of the potential next steps with regards to longer term follow-up of the trainee committee cohort, examines the wider implications of this work and considers an example of where the learning from this project has already successfully influenced the development of an award-winning leadership development programme.

8.3.1 Next steps – longer term follow-up

As highlighted in the sub-section above, a significant weakness of this study was that the 'end-point' (post-test) quantitative questionnaire detailed in chapter 6, and the post-intervention focus groups described in chapter 7, only took place around 12 months after the programme was initiated. These will therefore have only explored relatively early changes in knowledge, understanding and perspective of the participants. For any sort of formal analysis of longer term impact on individuals and groups of trainees in terms of behaviours, attitudes and achievements, the cohort, and ideally some of the participants who dropped out, would have to be revisited with further assessments a number of years down the line.

Whilst re-issuing the leadership self-assessment questionnaire to the trainee committee cohort for a third time would be fairly straightforward, in itself the conclusions that could be drawn from the results would be relatively limited as the breadth of confounding factors would have increased yet further, and there may be some questionnaire fatigue. The most interesting data would come from a qualitative approach, either through repeating the focus groups, or perhaps through conducting some semi-structured interviews with the original participants. This would enable the researchers to explore whether the leadership development intervention had any long-term influence on the participants in terms of their attitudes, perspectives and behaviours. A potential further source of data, which could be obtained through a number of different methods, could come from the observations and views of professional colleagues who are working alongside them. Although the logistics would be very difficult, as described in the sub-section above, it may also be possible to go back and re-assess the original comparator group.

8.3.2 Next steps – wider implications of this work

This study, which has focused on exploring whether trainee paediatricians can learn clinical leadership skills and capabilities, has a number of wider implications. Within the context of an extensive leadership literature that is relatively sparse on evidence for the value of specific interventions, it provides a mixed methods evaluation model that despite weaknesses could be adapted, improved and utilised in other studies. The combination and underlying rationale of the different statistical methods that were chosen to analyse the self-assessment data, and the use of focus groups to generate qualitative data, could also provide a useful framework for other researchers to adapt and use. In addition, with logistics often dictating the limitations of evaluative approaches to these sorts of leadership initiatives (to the extent that most are not evaluated at all) the potential for other studies to adapt and utilise the leadership abilities scale that was developed for this project could be of value.

Looking more broadly at the learning from the study, this work provides evidence to educational commissioners, training programme directors and other educational leaders of the value of a work-based team-focused approach to leadership learning. Although postgraduate paediatric trainees in London may not be representative of doctors, medical students and indeed other clinicians across all of the different levels of training and geographical locations, there are a number of ideas that have emerged from this study that might be applicable more widely. The need to focus learning both on the personal development of self-awareness and gaining perspectives of others, the building of team-working skills and finding opportunities for broader learning about the wider healthcare system are important areas of emphasis. These areas of learning align well with many of the competences described in the Medical Leadership Competency Framework (19) and associated Medical Leadership Curriculum (56). Gaining this learning through working together on a number of different improvement projects and initiatives as a trainee-led team, with consultant support and enabling where needed, appear to also be important factors.

Since 2010, in the UK, there has been increasing work by the Medical Royal Colleges, the General Medical Council and Postgraduate Deaneries to embed the leadership competencies illustrated within the Medical Leadership Competency Framework (MLCF) (19) into all postgraduate medical curricula and training programmes. Whilst this is an important step, the key challenge is in successfully creating a supportive learning environment and a strong network of dynamic trainers who can both support and enable trainees to develop their leadership capabilities through experiential learning. This project puts forward the argument that this development needs to happen throughout a trainee's journey within a training programme, rather than encouraging the historically accepted norm that attending a course in the weeks before a consultant interview will 'tick the leadership and management box'. It has also shown that the Medical Leadership Competency Framework (1) can provide a strong structure onto which a series of different leadership learning opportunities can hang.

In addition to offering a valuable opportunity to gain feedback and peer-review, the presentation and publication of this work is a key method of sharing the learning with the academic and education communities. An overview of the project, with the basic methodologies and results used in the pre-intervention phase, was published in a leading leadership journal (208) and the qualitative data from chapter 5 was presented as a poster at a regional conference (209). This generated feedback and reflection that was used to refine the analysis and discussion of the work. Many of the ideas and principles generated from this

thesis were built upon by the lead researcher (RK) in his co-authoring of the leadership book 'Medical Leadership – a practical guide for trainees and trainers' which was published in November 2011 (12). The project was highlighted in a leading leadership textbook (2), and the work has also stimulated a number of other different publications focused on different models of work-based leadership learning (207)(210)(211).

There is also some important learning that has been generated within this thesis that the lead researcher proposes submitting to peer-reviewed journals. One paper describes the overall mixed-methods approach to analysing the impact of the work-based leadership development programme detailed in this thesis. A second paper focuses on a wider analysis of the pre-intervention quantitative data described in chapter 4, using both the trainee committee cohort and comparator group data to draw conclusions around the learning needs of different levels of trainee. Finally, a third paper is being prepared which describes the different steps in the validation of the leadership abilities scale in the hope that it might be of value in other research.

8.3.3 Next steps – implementation of learning from this project into other initiatives

An important outcome of this project has been how the learning from it has already been directly used to develop other leadership initiatives. The most compelling example of this is with the development of 'Paired Learning', which is summarised in this sub-section. During the latter phases of this study it had occurred to the lead researcher (RK) that whilst participating in team-based leadership learning workshops and working together on meaningful projects (in this case to improve training, but in other examples may be about improving patient care) are effective ways to develop leadership capabilities, the addition of two other components could reap further benefits. At around the same time, the post-intervention 'end-point' focus groups had highlighted a fundamental shift in some participants in terms of an improved understanding of the perspectives of others, and there had also been a brief mention of the value of shadowing, or sitting-in on different experiences.

This emphasis on the impact of developing a perspective of others, led to reflection on the often toxic "them and us" relationships between managers and clinicians in the NHS. The vast majority of trainees in UK training programmes have never even said 'hello' to the managers in their department, let alone worked alongside them or sought to learn from them. There is increasing evidence that organisations who engender a strong culture of engagement between managers and clinicians provide safer, better quality care for patients (53)(212). This evidence, alongside the learning about principles of work-based leadership development from this study,

led to the design and implementation of a peer-learning leadership development programme called ‘Paired Learning’. This programme was led by the lead researcher (RK) alongside a senior management colleague. In the pilot year of this project (2010-11) seventeen pairs of managers and Specialist Registrar (SpR)-level clinicians participated in a 9-month long programme. The pairs used conversation, and in many cases shadowing (eg managers doing an on-call with the SpR, or the SpR attending a finance meeting), to learn about each other’s roles, responsibilities and perspectives. The pairs also worked together on a quality improvement project and were supported in their learning by five workshops which focused on the development of self-awareness, team-working, understanding the system, tools for change and patient safety. The four components of ‘Paired Learning’ are illustrated in Figure 8.1.

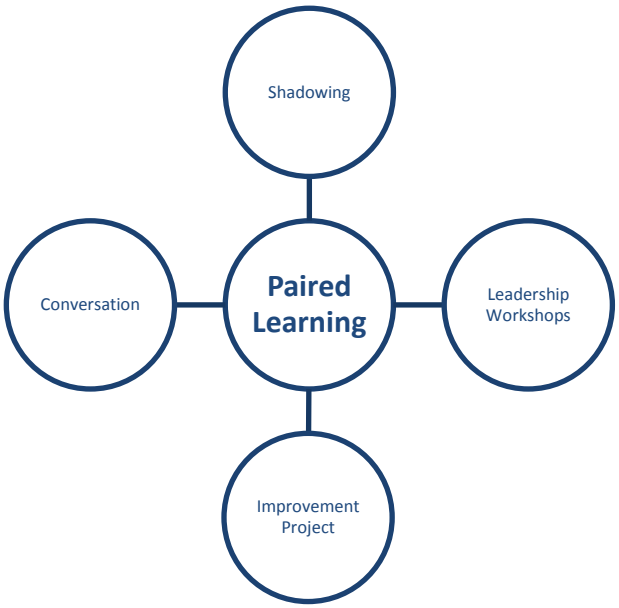


Figure 8.1 Outline of the four components of ‘Paired Learning’

Based on the lead researcher’s experiences and learning from this thesis, a mixed methods approach to the evaluation of ‘Paired Learning’ was taken. This consisted of pre- and post-intervention self-assessment questionnaires (which used the abilities scale validated in this project) alongside semi-structured interviews of the participants at the end of the programme. The evaluation demonstrated that the programme significantly increased participants’ preparedness for leadership across a wide range of dimensions. The qualitative analysis demonstrated that the co-development of managers and doctors had a powerful impact on the personal learning, attitudes and behaviour of participants. In addition there were a number of demonstrable wider organisational benefits, which resulted in improvements to patient care through the collaborative work done within the programme (206)(120). The pilot project won a

major regional education award, was runner-up in a national competition and continues to be adopted by increasing numbers of organisations from across the NHS.

8.4 Personal learning from this work

Undertaking this work alongside developing my clinical and educational career has been challenging but extremely rewarding, and has provided me with some valuable experiences and a huge amount of personal learning.

The project has allowed me to develop my interest in work-based learning, and in particular how doctors and other professionals can develop leadership and team-working capabilities. This has given me the skills and confidence to design and implement other leadership programmes and learning opportunities which have now benefited a large number of different participants from across the health service.

My medical education training and experience prior to this thesis had provided me with a broad comprehension of different evaluative methodologies. However, this project has given me a much deeper understanding of mixed methods research. The detailed discussions that took place during the oral viva of this thesis, and the reading undertaken since, have strengthened this understanding significantly. Most revealing has been to see how a combination of different research methods, when sitting within a well designed experiment, can overcome the classic difficulties of educational research with small numbers of participants and multiple variables that are near impossible to measure. The development of skills in methodological design, data collection, framework analysis and statistical analysis has been extremely useful. These have already been utilised in designing and implementing the evaluation of the 'Paired Learning' programme described in sub-section 8.3.3 above.

In addition to learning new research skills the experiences from this work have allowed me to reflect on two important concepts that underpin any learning and development. The first of these concepts is around the value of seeking many different ideas and opinions in order to shape an intervention accordingly. The opportunity to share and discuss ideas with colleagues who have different perspectives and experiences is extremely valuable and can hugely influence the development of different projects or initiatives. The second concept is of the value of synergy; where different projects have similar or over-lapping aims, processes or outcomes there is great potential for cross-over, combining of resources and sharing of

learning. Both of these approaches encourage the output of high quality initiatives that have the best possible chance of making a significant impact.

8.5 Future research

In this final section, three key areas for future work need to be emphasised. The main area for research directly linked to this work would be to establish a longitudinal study that followed up the original trainee committee participants over a 3,5 and perhaps 10 year period. The aim would be to explore whether participants' experiences of the trainee committee and the work-based leadership development initiative that was built around it had any long term impact on their development and career progress. It would be particularly interesting to use qualitative methodologies to explore some of the behavioural and organisational areas of change highlighted by different evaluation methodologies (89)(133)(64).

The 'Paired Learning' work, described in subsection 8.3.3 above, has been an exciting development which has arisen from this project. Publishing the longer term follow-up data from this work is a key priority, with focus on exploring both personal learning and also any organisational and patient outcomes that have arisen from within the programme. Any further evidence of benefits here would significantly help with the sustainability and spread of the initiative across a number of different organisations.

Perhaps, above all, is the importance that the education and organisational development (OD) communities recognise that in order to understand the impact and value of different leadership development approaches, programmes need to be underpinned by methodologically strong, well-planned and resourced evaluation strategies. Only when meaningful evaluation is embedded into the culture of all leadership development work within healthcare will we be able to develop programmes that are truly fit for purpose.

References

1. Northouse PG. Leadership: theory and practice. Thousand Oaks: SAGE; 2013.

2. Spurgeon P, Clark J, Ham C. Medical Leadership: From the Dark Side to Centre Stage. 1st ed. Radcliffe Publishing Ltd; 2011.

3. Kotter JP. Force For Change: How Leadership Differs from Management. Free Press; 1990.

4. Pervin LA. A Critical Analysis of Current Trait Theory. Psychological Inquiry. 1994 Apr 1;5(2):103–13.

5. Judge TA, Colbert AE, Ilies R. Intelligence and leadership: a quantitative review and test of theoretical propositions. J Appl Psychol. 2004 Jun;89(3):542–52.

6. Alimo-Metcalfe B, Alban-Metcalfe RJ. The development of a new Transformational Leadership Questionnaire. Journal of Occupational and Organizational Psychology. 2001;74(1):1–27.

7. Gill R. Theory and Practice of Leadership. SAGE; 2011.

8. Bass BM, Bass R. The Bass handbook of leadership: theory, research, and application. New York: Free Press; 2008.

9. Tannenbaum R, Schmidt WH. How To Choose a Leadership Pattern. Harvard Business Review. 1973 Jan;

10. Hartley J, Benington J. Leadership for healthcare. Bristol; Portland, OR: Policy Press; 2010.

11. Blake RR, Mouton JS. The managerial grid: key orientations for achieving production through people. Gulf Pub. Co.; 1964.

12. Spurgeon P, Klaber R. Medical Leadership: A practical guide for trainees & tutors. 1st edition. London: BPP Learning Media; 2011.

13. Cartwright D, Zander AF. Group dynamics: research and theory. Harper & Row; 1968.

14. McGregor D. The human side of enterprise. McGraw-Hill; 1960.

15. Korman AK. "Consideration,""Initiating Structure," and Organizational Criteria- a Review. Personnel Psychology. 1966 Dec;19(4):349–61.

16. Kerr S, Schriesheim CA, Murphy CJ, Stogdill RM. Toward a contingency theory of leadership based upon the consideration and initiating structure literature. Organizational Behavior and Human Performance. 1974 Aug;12(1):62–82.

17. Derue DS, Nahrgang JD, Wellman NED, Humphrey SE. Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. Personnel Psychology. 2011;64(1):7–52.

18. Goffee R, Jones G. Why should anyone be led by you? Harv Bus Rev. 2000 Oct;78(5):62–70, 198.

19. Academy of Medical Royal Colleges and NHS Institute for Innovation and Improvement. Medical Leadership Competency Framework Coventry: NHS Institute for Innovation and Improvement; 2008. Available from: www.institute.nhs.uk/mlcf [accessed 23 05 15]

20. Burns JM. Leadership. Harper & Row; 1978.

21. Bass BM, Avolio BJ. Improving Organizational Effectiveness Through Transformational Leadership. SAGE; 1994.

22. Grint K. Management or leadership? J Health Serv Res Policy. 2002 Oct;7(4):248–51.

23. Swanwick T, McKimm J. What is clinical leadership…and why is it important? The Clinical Teacher. 2011;8(1):22–6.

24. Gronn P. The Making of Educational Leaders. Cassell; 1999.

25. Mintzberg H. The structuring of organizations: a synthesis of the research. Prentice-Hall; 1979.

26. Kellerman B. Followership: How Followers are Creating Change and Changing Leaders. Harvard Business School Press; 2008.

27. Badrick T, Preston A. Influences on the implementation of TQM in health care organizations: professional bureaucracies, ownership and complexity. Aust Health Rev. 2001;24(1):166–75.

28. Batalden PB, Nelson EC, Mohr JJ, Godfrey MM, Huber TP, Kosnik L, et al. Microsystems in health care: Part 5. How leaders are leading. Jt Comm J Qual Saf. 2003 Jun;29(6):297–308.

29. Conger JA. Leadership: The Art of Empowering Others. The Academy of Management Executive (1987-1989). 1989 Feb;3(1):17–24.

30. Pearce CL, Conger J. Shared Leadership: Reframing the Hows and Whys of Leadership. Sage Publications Thousand Oaks, Calif; 2002.

31. Pearce CL, Manz CC, Sims Jr. HP. Where Do We Go From Here?:: Is Shared Leadership the Key to Team Success? Organizational Dynamics. 2009 Jul;38(3):234–8.

32. West M, Amit K, Loewenthal L, Eckert R, West T, Lee A. Leadership and leadership development in health care. London: The Kings Fund; 2015. http://www.kingsfund.org.uk/publications/leadership-and-leadership-development-health-care [accessed 23 05 15]

33. Konu A, Viitanen E. Shared leadership in Finnish social and health care. Leadership in Health Services. 2008 Feb;21(1):28–40.

34. Ulhøi JP, Müller S. Mapping the Landscape of Shared Leadership: A Review and Synthesis. International Journal of Leadership Studies. 2014;8(2):66–87.

35. The NHS Information Centre. NHS Staff 2000 - 2010 Overview. http://www.ic.nhs.uk/statistics-and-data-collections/workforce/nhs-staff-numbers/nhs-staff-2000--2010-overview [accessed 18 05 15]

36. Mountford J, Webb C. When clinicians lead. The McKinsey Quarterly. 2009 Feb.

37. Mountford J, Webb C. Clinical Leadership - unlocking high performance in healthcare. McKinsey & Company; 2008.

38. Stanton E, Lemer C, Mountford J. Clinical leadership: bridging the divide. London: Quay Books; 2010.

39. Storey J, Holti R. Possibilities and pitfalls for clinical leadership in improving service quality, innovation and productivity. London, UK: NIHR Service Delivery and Organisation programme; 2013. http://oro.open.ac.uk/36270/1/SDO_FR_09-1001-22_V05.pdf [accessed 23 05 15]

40. Berry LL, Seltman. Management lessons from Mayo Clinic: inside one of the world's most admired service organizations. New York: McGraw-Hill; 2008.

41. Crosson FJ. Improving the doctor-manager relationship. Kaiser Permanente: a propensity for partnership. BMJ. 2003 Mar;326(7390):654.

42. Morris S, Hunter RM, Ramsay AIG, Boaden R, McKevitt C, Perry C, et al. Impact of centralising acute stroke services in English metropolitan areas on mortality and length of hospital stay: difference-in-differences analysis. BMJ. 2014 Aug;349:g4757.

43. Darzi A. High Quality Care for All: NHS Next Stage Review Final Report. London: Department of Health; 2008.

44. Tooke J. Aspiring to Excellence: Findings and recommendations of the Independent Inquiry into Modernising Medical Careers. London: Aldridge Press; 2008.

45. Ham C, Dickinson H. Engaging Doctors in Leadership: What we can learn from international experience and research evidence? Coventry, UK: NHS Institute for Innovation and Improvement; 2008.

46. Leadership and management for all doctors. London, UK: General Medical Council; 2012. http://www.gmc-uk.org/Leadership_and_management_for_all_doctors___English_0914.pdf_48903400.pdf [accessed 23 05 15]

47. Bohmer RMJ. Leading Clinicians and Clinicians Leading. New England Journal of Medicine. 2013 Apr;368(16):1468–70.

48. Gilbert A, Hockey P, Vaithianathan R, Curzen N, Lees P. Perceptions of junior doctors in the NHS about their training: results of a regional questionnaire. BMJ Qual Saf. 2012 Mar;21(3):234–8.

49. Batalden PB, Davidoff F. What is "quality improvement" and how can it transform healthcare? Qual Saf Health Care. 2007 Feb;16(1):2–3.

50. Dickinson H, Ham C, Snelling I, Spurgeon P. Are We There Yet? Models of Medical Leadership and their effectiveness: An Exploratory Study. NIHR Service Delivery and Organisation programme; 2013. http://www.nets.nihr.ac.uk/__data/assets/pdf_file/0007/85066/FR-08-1808-236.pdf [accessed 23 05 15]

51. Case Studies Archives. NHS Leadership Academy. 2015. http://www.leadershipacademy.nhs.uk/case-studies/ [accessed 23 05 15]

52. Leadership and Management Standards for Medical Professionals. London, UK: Faculty of Medical Leadership and Management; 2015. https://www.fmlm.ac.uk/professional-

development/accreditation-and-standards/the-leadership-and-management-standards-for [accessed 23 05 15]

53. Spurgeon P, Mazelan PM, Barwell F. Medical engagement: a crucial underpinning to organizational performance. Health Serv Manage Res. 2011 Aug;24(3):114–20.

54. NHS Institute for Innovation and Improvement. NHS Leadership Qualities Framework. 2006.

55. Tomorrow's Doctors. London, UK: General Medical Council; 2009. http://www.gmc-uk.org/Tomorrow_s_Doctors_0414.pdf_48905759.pdf [accessed 23 05 15]

56. Medical Leadership Curriculum. Coventry, UK: NHS Institute for Innovation and Improvement; 2010. https://www.institute.nhs.uk/images/documents/Leadership/Medical%20Leadership%20Curriculum.pdf [accessed 23 05 15]

57. RCPCH. Facing the Future: Standards for Paediatric Services. London: RCPCH; 2010.

58. Facing the Future: Together for child health. London, UK: RCPCH; 2015 http://www.rcpch.ac.uk/sites/default/files/page/Facing%20the%20Future%20Together%20for%20Child%20Health%20final%20web%20version.pdf [accessed 23 05 15]

59. Easen P, Atkins M, Dyson A. Inter-professional collaboration and conceptualisations of practice. Children & Society. 2000 Nov;14(5):355–67.

60. Doyle J. Barriers and facilitators of multidisciplinary team working: a review. Paediatric Nursing. 2008 Mar;20(2):26–9.

61. Jackson B, Parry KW. A very short, fairly interesting and reasonably cheap book about studying leadership. Los Angeles; London: SAGE; 2011.

62. Zaleznik A. Managers and leaders: are they different? Clin Leadersh Manag Rev. 2004 Jun;18(3):171–7.

63. Chambers R. How to succeed as a leader. Oxford; New York: Radcliffe Pub.; 2007.

64. Wisdom JP, Cavaleri MA, Onwuegbuzie AJ, Green CA. Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. Health Serv Res. 2012 Apr;47(2):721–45.

65. Maudsley G. Mixing it but not mixed-up: mixed methods research in medical education (a critical narrative review). Med Teach. 2011;33(2):e92–104.

66. Schifferdecker KE, Reed VA. Using mixed methods research in medical education: basic guidelines for researchers. Med Educ. 2009 Jul;43(7):637–44.

67. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. BMJ. 2010 Sep;341:c4587–c4587.

68. Doyal Report. The Ethical Governance and Regulation of Student Projects: Working Group on Ethical Review of Student Research in the NHS. London: Department of Health; 2005.

69. Pugsley L, Dornan T. Using a sledgehammer to crack a nut: clinical ethics review and medical education research projects. Medical Education. 2007 Aug;41(8):726–8.

70. Research Ethics Proforma. School of Postgraduate Medical and Dental Education, Cardiff University, Cardiff, UK; 2007.

71. Eva KW. Research ethics requirements for Medical Education. Medical Education. 2009 Mar;43(3):194–5.

72. Cooper RA, Tauber AI. Values and ethics: a collection of curricular reforms for a new generation of physicians. Acad Med. 2007 Apr;82(4):321–3.

73. Haig A, Dozier M. BEME Guide no 3: systematic searching for evidence in medical education--Part 1: Sources of information. Med Teach. 2003 Jul;25(4):352–63.

74. Cummings G, Lee H, Macgregor T, Davey M, Wong C, Paul L, et al. Factors contributing to nursing leadership: a systematic review. J Health Serv Res Policy. 2008;13(4):240–8.

75. Auger CP. Information Sources in Grey Literature. Bowker-Saur; 1994.

76. Jacobs S, Kelleen M. Research Guides. Health (Nursing, Medicine, Allied Health). Grey Literature.2015. Available from: http://guides.nyu.edu/greylithealth [accessed 23 05 15]

77. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions [In: The Cochrane Library, Issue 4, 2006]. Chichester, West Sussex, UK: John Wiley & Sons, Ltd; 2006.

78. Haig A, Dozier M. BEME guide no. 3: systematic searching for evidence in medical education-part 2: constructing searches. Med Teach. 2003 Sep;25(5):463–84.

79. Fitzpatrick RB. TIMELIT: topics in medical education database. Med Ref Serv Q. 2009;28(1):56–64.

80. Chakraborti C, Boonyasai RT, Wright SM, Kern DE. A systematic review of teamwork training interventions in medical student and resident education. J Gen Intern Med. 2008 Jun;23(6):846–53.

81. Steinert Y, Naismith L, Mann K. Faculty development initiatives designed to promote leadership in medical education. A BEME systematic review: BEME Guide No. 19. Med Teach. 2012;34(6):483–503.

82. Duda RB. The impact of a leadership development course for junior academic medical faculty. Leadership in Health Services. 2008 Jul;21(3):200–7.

83. Osborn LM, DeWitt T. The HRSA-APA Faculty Development Scholars Program: Executive Leadership Track. Ambulatory Pediatrics. 2004 Jan;4(1):98–102.

84. Juriza I, Ruzzana I, Harlina H, Rohaizak R. Outdoor Camps Experiential Learning Activities for Teamwork and Leadership among Medical Students. Procedia - Social and Behavioral Sciences. 2011;18:622–5.

85. Evans DV, Egnew TR. Outdoor-based leadership training and group development of family practice interns. Fam Med. 1997 Aug;29(7):471–6.

86. Stoller JK, Rose M, Lee R, Dolgan C, Hoogwerf BJ. Teambuilding and leadership training in an internal medicine residency training program. J Gen Intern Med. 2004 Jun;19(6):692–7.

87. Smith KL, Petersen DJ, Soriano R, Friedman E, Bensinger LD. Training Tomorrow's Teachers Today: a national medical student teaching and leadership retreat. Med Teach. 2007 May;29(4):328–34.

88. Dobson C., Cockson J., Allgar V., McKendree J. Leadership training in the undergraduate medical curriculum. Education for Primary Care. 2008;19(5):526–9.

89. Kirkpatrick DL. Evaluating training programs: a collection of articles from the Journal of the American Society for Training and Development. American Society for Training and Development; 1975.

90. Leslie LK, Miotto MB, Liu GC, Ziemnik S, Cabrera AG, Calma S, et al. Training young pediatricians as leaders for the 21st century. Pediatrics. 2005 Mar;115(3):765–73.

91. Steinert Y, Nasmith L, Daigle N. Executive skills for medical faculty: a workshop description and evaluation. Med Teach. 2003 Nov;25(6):666–8.

92. Hemmer PR, Karon BS, Hernandez JS, Cuthbert C, Fidler ME, Tazelaar HD. Leadership and management training for residents and fellows: a curriculum for future medical directors. Archives of Pathology & Laboratory Medicine. 2007;131(4):610–4.

93. McAlearney AS, Fisher D, Heiser K, Robbins D, Kelleher K. Developing effective physician leaders: changing cultures and transforming organizations. Hosp Top. 2005;83(2):11–8.

94. Morahan P, Kasperbauer D, McDade S, Aschenbrener C, Triolo P, Monteleone P, et al. Training future leaders of academic medicine: internal programs at three academic health centers. Academic Medicine. 1998;73(11):1159–68.

95. Dannels SA, Yamagata H, McDade SA, Chuang Y-C, Gleason KA, McLaughlin JM, et al. Evaluating a Leadership Program: A Comparative, Longitudinal Study to Assess the Impact of the Executive Leadership in Academic Medicine (ELAM) Program for Women. Academic Medicine. 2008 May;83(5):488–95.

96. Shapiro MJ, Morey JC, Small SD, Langford V, Kaylor CJ, Jagminas L, et al. Simulation based teamwork training for emergency department staff: does it improve clinical team performance when added to an existing didactic teamwork curriculum? Qual Saf Health Care. 2004 Dec;13(6):417–21.

97. Ostergaard H, Ostergaard D, Lippert A. Implementation of team training in medical education in Denmark. Qual Saf Health Care. 2004 Oct;13(Suppl 1):i91–5.

98. Dienst ER, Byl N. Evaluation of an educational program in health care teams. J Community Health. 1981;6(4):282–98.

99. Cooper H, Spencer-Dawe E, McLean E. Beginning the process of teamwork: design, implementation and evaluation of an inter-professional education intervention for first year undergraduate students. J Interprof Care. 2005 Oct;19(5):492–508.

100. McAlearney AS. Using leadership development programs to improve quality and efficiency in healthcare. J Healthc Manag. 2008 Oct;53(5):319–31; discussion 331–2.

101. Sawer P, Ball J. NHS chief tells trusts to make £20bn savings. The Daily Telegraph. 2009 Jun 13; http://www.telegraph.co.uk/health/healthnews/5524693/NHS-chief-tells-trusts-to-make-20bn-savings.html [accessed 13 09 2012]

102. Beecham B., Dammers J., Van Zwanenberg T. Leadership coaching for general practitioners. Education for Primary Care. 2004;15(4):579–83.

103. Myers I. The Myers-Briggs Type Indicator. CA: Consulting Psychologists Press; 1962.

104. Shope TC, Frohna JG, Frohna AZ. Using the Myers-Briggs Type Indicator (MBTI) in the teaching of leadership skills. Med Educ. 2000 Nov;34(11):956.

105. Wahlstrom O, Sanden I. Multiprofessional training ward at Linköping University, early experience. Education for Health. 1998;11:225–31.

106. Hill F, Stephens C. Building leadership capacity in medical education: developing the potential of course coordinators. Medical Teacher. 2005 Jan;27(2):145–9.

107. Crites GE, Ebert JR, Schuster RJ. Beyond the dual degree: development of a five-year program in leadership for medical undergraduates. Academic Medicine. 2008;83(1):52–8.

108. Schwartz RW, Pogge CR, Gillis SA, Holsinger JW. Programs for the development of physician leaders: a curricular process in its infancy. Acad Med. 2000 Feb;75(2):133–40.

109. Kolb D. Experiential Learning: Experience as the Source of Learning and Development. Englewood Cliffs, NJ: Prentice Hall; 1984.

110. Gowan I. A leadership programme. BMJ Careers. 2010 Apr;

111. Coltart CE, Cheung R, Ardolino A, Bray B, Rocos B, Bailey A, et al. Leadership development for early career doctors. The Lancet. 2012 May;379(9828):1847–9.

112. Dickinson H, Ham C. Engaging Doctors in Leadership: Review of the Literature. Birmingham: Health Services Management Centre; 2008.

113. Stoll L. Leadership Learning: Designing a Connected Strategy. Jolimont, Victoria: Incorporated Association of Registered Teachers of Victoria; 2004.

114. Hallinger P. Reshaping the Landscape of School Leadership Development. Taylor & Francis; 2003.

115. General Paediatric Curriculum. London, UK: RCPCH; 2010.

116. Levenson R, Atkinson S, Shepherd S. The 21st Century Doctor: Understanding the doctors of tomorrow. London, UK: The Kings Fund; 2010. http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/Understanding-doctors-tomorrow-levenson-atkinson-shepherd-oct10.pdf [accessed 21 02 2014]

117. Cox JF, Pearce CL, Perry ML. Shared leadership: Reframing the hows and whys of leadership. 2003;

118. A Reference Guide for Postgraduate Specialty Training in the UK. COPMeD; 2014 May.

119. Stoll L, Foster-Turner J, Glenn M. Mind Shift: an evaluation of the NHS London "Darzi" Fellowships in Clinical Leadership Programme. London, UK: Institute of Education; 2010.

120. Klaber RE, Lee J. Clinical leadership and management in the NHS – Paired Learning. J R Soc Med. 2011 Nov;104(11):436–436.

121. Myers IBMMH; Q, Myers, Isabel Briggs; McCaulley Mary H.; Quenk, Naomi L.; Hammer, Allen L. MBTI Manual (A guide to the development and use of the Myers Briggs type indicator). Consulting Psychologists Press; 3rd ed edition; 1998.

122. Belbin RM. Management Teams: Why They Succeed Or Fail. Butterworth-Heinemann; 1981.

123. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. Medical Teacher. 2006 Jan;28(7):e185–91.

124. GMP Framework for appraisal and revalidation. Available from: http://www.gmc-uk.org/doctors/revalidation/revalidation_gmp_framework.asp [accessed 19 08 2012]

125. Campbell JL, Richards SH, Dickens A, Greco M, Narayanan A, Brearley S. Assessing the professional performance of UK doctors: an evaluation of the utility of the General Medical Council patient and colleague questionnaires. Qual Saf Health Care. 2008 Jun;17(3):187–93.

126. Yardley S, Dornan T. Kirkpatrick's levels and education "evidence." Med Educ. 2012 Jan;46(1):97–106.

127. Coldwell M, Simkins T. Level models of continuing professional development evaluation: a grounded review and critique. Professional Development in Education. 2011 Feb;37(1):143–57.

128. Alliger GM, Janak EA. Kirkpatrick's Levels of Training Criteria: Thirty Years Later. Personnel Psychology. 1989;42(2):331–42.

129. Alliger GM, Tannenbaum SI, Bennett W, Traver H, Shotland A. A Meta-Analysis of the Relations Among Training Criteria. Personnel Psychology. 1997;50(2):341–58.

130. Bates R. A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. Evaluation and Program Planning. 2004;(3):341–7.

131. Stufflebeam DL, American Evaluation Association. Evaluation models. San Francisco: Jossey-Bass; 2001.

132. Stufflebeam DL, Coryn CLS. Evaluation Theory, Models, and Applications. John Wiley & Sons; 2014.

133. Kaufman R, Keller J, Watkins R. What works and what doesn't: Evaluation beyond kirkpatrick. Nonprofit Management Leadership. 1996 Feb;35(2):8–12.

134. Parker K, Burrows G, Nash H, Rosenblum ND. Going beyond Kirkpatrick in evaluating a clinician scientist program: it's not "if it works" but "how it works." Acad Med. 2011 Nov;86(11):1389–96.

135. Frye AW, Hemmer PA. Program evaluation models and related theories: AMEE guide no. 67. Med Teach. 2012;34(5):e288–99.

136. Frechtling J. Logic Modeling Methods in Program Evaluation [Internet]. Jossey-Bass: Wiley; 2007. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0787981966.html [accessed 28 05 2015]

137. Levin-Rozalis M. Evaluation and research differences and similarities. The Canadian Journal of Program Evaluation. 2003;18(2):1–31.

138. Fain JA. Is There a Difference Between Evaluation and Research? The Diabetes Educator. 2005 Mar;31(2):150–5.

139. Scriven M. Evaluation Thesaurus. SAGE; 1991.

140. Smith NL, Brandon PR. Fundamental Issues in Evaluation. Guilford Press; 2008.

141. Patton MQ. Utilization-focused evaluation. Thousand Oaks, Calif.: Sage Publications; 2008.

142. Lingard L, Albert M, Levinson W. Grounded theory, mixed methods, and action research. BMJ. 2008 Aug;337:a567.

143. Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine; 1967.

144. Glaser BG. Basics of Grounded Theory Analysis: Emergence Vs. Forcing. Sociology Press; 1992.

145. Strauss AL. Qualitative Analysis for Social Scientists. Cambridge University Press; 1987.

146. Campbell DT, Stanley JC, Gage NL. Experimental and quasi-experimental designs for research. Boston: Houghton Mifflin; 1963.

147. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. SAGE Publications; 2013.

148. Greene JC, Caracelli VJ, Graham WF. Toward a Conceptual Framework for Mixed-Method Evaluation Designs. Educational Evaluation and Policy Analysis. 1989 Sep;11(3):255–74.

149. Creswell JW, Plano Clark VL. Designing and conducting mixed methods research. Los Angeles: SAGE Publications; 2011.

150. Greenhalgh T, Collard A, Begum N. Sharing stories: complex intervention for diabetes education in minority ethnic groups who do not speak English. BMJ. 2005 Mar;330(7492):628.

151. Davies H, Archer J, Bateman A, Dewar S, Crossley J, Grant J, et al. Specialty-specific multi-source feedback: assuring validity, informing training. Med Educ. 2008 Oct;42(10):1014–20.

152. Archer J, McGraw M, Davies H. Assuring validity of multisource feedback in a national programme. Arch Dis Child. 2010 May;95(5):330–5.

153. Archer JC, McAvoy P. Factors that might undermine the validity of patient and multi-source feedback. Med Educ. 2011 Sep;45(9):886–93.

154. Brown JM, Lowe K, Fillingham J, Murphy PN, Bamforth M, Shaw NJB. An investigation into the use of multi-source feedback (MSF) as a work-based assessment tool. Med Teach. 2014 Nov;36(11):997–1004.

155. Al Ansari A, Donnon T, Al Khalifa K, Darwish A, Violato C. The construct and criterion validity of the multi-source feedback process to assess physician performance: a meta-analysis. Adv Med Educ Pract. 2014;5:39–51.

156. Donnon T, Al Ansari A, Al Alawi S, Violato C. The reliability, validity, and feasibility of multisource feedback physician assessment: a systematic review. Acad Med. 2014 Mar;89(3):511–6.

157. NHS Leadership Academy. Leadership Framework Self Assessment Tool. Coventry: NHS Institute for Innovation and Improvement; 2012. http://www.leadershipacademy.nhs.uk/develop-your-leadership-skills/leadership-framework/supporting-tools-and-documents/self-assessment-tools [accessed 26 06 2012]

158. NHS Leadership Academy. Leadership Framework. Coventry: NHS Institute for Innovation and Improvement; 2011.

159. Robert G, Cornwell J, Locock L, Purushotham A, Sturmey G, Gager M. Patients and staff as codesigners of healthcare services. BMJ. 2015 Feb;350:g7714.

160. CUREE Evaluation of Learning Away: Hypothesis 7: Student leadership, co-design and facilitation of learning. London: Paul Hamlyn Foundation; 2012.

161. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990 Sep;65(9 Suppl):S63–7.

162. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). Medical Education. 2009 Jan;13(1):39–54.

163. Pell G, Fuller R, Homer M, Roberts T. How to measure the quality of the OSCE: A review of metrics - AMEE guide no. 49. Med Teach. 2010;32(10):802–11.

164. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA. 2006 Sep;296(9):1094–102.

165. Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. Adv Health Sci Educ Theory Pract. 2002;7(1):63–80.

166. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005 Oct;80(10 Suppl):S46–54.

167. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol. 1999 Dec;77(6):1121–34.

168. Colthart I, Bagnall G, Evans A, Allbutt H, Haig A, Illing J, et al. The effectiveness of self-assessment on the identification of learner needs, learner activity, and impact on clinical practice: BEME Guide no. 10. Med Teach. 2008;30(2):124–45.

169. Kouzes JM, Posner BZ. Leadership practices inventory: leadership development planner. Pfeiffer; 2003.

170. Martin JS, Mccormack B, Fitzsimons D, Spirig R. Evaluation of a clinical leadership programme for nurse leaders. Journal of Nursing Management. 2011 Jun;20(1):72–80.

171. Alexander CS, Becker HJ. The Use of Vignettes in Survey Research. Public Opin Q. 1978 Mar;42(1):93–104.

172. Finch J. The vignette technique in survey research. Sociology. 1987;21:105–14.

173. Barter C, Renold E. “I wanna tell you a story”: the use of vignettes in qualitative research. International Journal of Social Research Methodology. 2000;3(4):307–23.

174. Van Soest A, Delaney L, Harmon C, Kapteyn A, Smith JP. Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2011;174(3):575–95.

175. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. JAMA. 2000 Apr;283(13):1715–22.

176. Veloski J, Tai S, Evans AS, Nash DB. Clinical vignette-based surveys: a tool for assessing physician practice variation. Am J Med Qual. 2005 Jun;20(3):151–7.

177. Hamilton P, Spurgeon P, Clark J, Dent J, Armit K. Engaging Doctors: Can Doctors Influence Organisational Performance? Coventry, UK: NHS Institute for Innovation and Improvement; 2008.

178. Osgood CE. The Measurement of Meaning. University of Illinois Press; 1957.

179. Oppenheim AN. Questionnaire Design, Interviewing and Attitude Measurement. Continuum; 1992.

180. Jamieson S. Likert scales: how to (ab)use them. Med Educ. 2004 Dec;38(12):1217–8.

181. Pell G. Use and misuse of Likert scales. Medical Education. 2005;39(9):970–970.

182. Carifio J, Perla RJ. Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. Journal of Social Sciences. 2007;3(3):106–16.

183. Dreyfus HL, Dreyfus SE. Mind Over Machine: The Power of Human Intuition and Expertise in the Computer Era. Oxford: Basil Blackwell; 1986.

184. Stewart J, O’Halloran C, Barton JR, Singleton SJ, Harrigan P, Spencer J. Clarifying the concepts of confidence and competence to produce appropriate self-evaluation measurement scales. Med Educ. 2000 Nov;34(11):903–9.

185. Iglesias C, Torgerson D. Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire. J Health Serv Res Policy. 2000 Oct;5(4):219–21.

186. Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, et al. Increasing response rates to postal questionnaires: systematic review. BMJ. 2002 May;324(7347):1183.

187. Hoffman SC, Burke AE, Helzlsouer KJ, Comstock GW. Controlled trial of the effect of length, incentives, and follow-up techniques on response to a mailed questionnaire. Am J Epidemiol. 1998 Nov;148(10):1007–11.

188. Hardigan PC, Succar CT, Fleisher JM. An analysis of response rate and economic costs between mail and web-based surveys among practicing dentists: a randomized trial. J Community Health. 2012 Apr;37(2):383–94.

189. Sedgwick P. Multiple significance tests: the Bonferroni correction. BMJ. 2012 Jan;344:e509–e509.

190. Wetzel AP. Factor analysis methods and validity evidence: a review of instrument development across the medical education continuum. Acad Med. 2012;87(8):1060–9.

191. D. L. Morgan. Focus Groups as Qualitative Research. London: Sage; 1988.

192. Barbour RS. Making sense of focus groups. Medical Education. 2005;39(7):742–50.

193. Krueger RA. Focus groups: a practical guide for applied research. Sage Pub; 1994.

194. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. Analysing Qualitative Data [A Bryman and RG Burgess, editors]. London: Routledge; 1994. p. 173–94.

195. Burrows D, Kendall S. Focus groups: what are they and how can they be used in nursing and health care research? Social Sciences in Health. 1997;3:244–53.

196. Pope C, Ziebland S, Mays N. Qualitative research in health care: Analysing qualitative data. BMJ. 2000 Jan;320(7227):114–6.

197. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. Sage Pub; 2000.

198. Rabiee F. Focus-Group Interview and Data Analysis. Proceedings of the Nutrition Society. 2004;63(04):655–60.

199. Moran-Ellis J, Alexander VD, Cronin A, Dickinson M, Fielding J, Sleney J, et al. Triangulation and integration: processes, claims and implications. Qualitative Research. 2006 Feb;6(1):45–59.

200. Green J, Thorogood N. Qualitative Methods in Health Research. Sage Pub; 2004.

201. Dicicco-Bloom B, Crabtree BF. The qualitative research interview. Med Educ. 2006 Apr;40(4):314–21.

202. Tariq S, Woodman J. Using mixed methods in health research. JRSM Short Rep. 2013 May;4(6).

203. Wheeldon J, Ahlberg MK. Visualizing Social Science Research: Maps, Methods, & Meaning. SAGE Publications; 2011.

204. Akobeng AK. Understanding randomised controlled trials. Arch Dis Child. 2005 Aug;90(8):840–4.

205. Hutchinson D, Styles B. A guide to running randomised controlled trials for educational researchers. Slough: National Foundation for Educational Research; 2011. Available from: http://www.nfer.ac.uk/nfer/publications/RCT01/RCT01.pdf [accessed 05 10 2012]

206. Klaber B, Lee J, Abraham R, Smith L, Lemer C. Paired Learning: a peer-learning leadership development initiative for managers and clinicians in the NHS. UK: NHS London; 2012 Feb.

207. Klaber B, Lee J, Abraham R, Lemer C, Smith L. Leadership: Two's company. Health Serv J. 2011 Sep;121(6275):21–3.

208. Klaber R, Roueche A, Hodgkinson R, Cass H. A structured approach to planning a work-based leadership development programme for doctors in training. The International Journal of Clinical Leadership. 2009;16(3):121–9.

209. Klaber RE, Hodgkinson R, Roueché A, Cass H. Leadership development in paediatrics – the trainee view. Poster presented at: UCLP Child Health Symposium; 2009 Sep; Institute of Child Health.

210. Roland D, Warren O, Klaber R. Engaging with leadership learning in the workplace. The Lancet. 2012 Aug;380(9841):563.

211. Spurgeon P, Klaber R. Becoming a better medical leader. BMJ Careers. 2012 Jan;

212. Spurgeon P, Barwell F, Mazelan P, Clark J. The Link Between Medical Engagement and Organisational Performance. Coventry, UK: NHS Institute for Innovation and Improvement; 2009.

Appendices

1.1 Ethics Committee approval letter

3.1 Trainee Committee Application Information Sheet

3.2 Leadership 360 participant information sheet

3.3 Leadership 360 invitation to participate

3.4 Printed version of full set of Leadership 360 questions

3.5 Thematic analysis of trainee committee application forms

4.1 Pre-intervention self-assessment questionnaire

4.2 Self-assessment ability score data (pre- and post-intervention)

5.1 Focus groups participant information sheet

5.2 Focus groups consent form

Appendix 3.1 Trainee Committee Application Information Sheet

What is the purpose of having a trainee committee?

Under the old system of Specialist Training Committees (STCs) there was a single trainee representative from North and South Thames who attended meetings. Despite the best efforts of the individuals involved, broad representation of the full cross-section of trainees from throughout the region was impossible. This new committee, which will consist of around 20 trainees, will help considerably with this. The trainee committee will be asked to consider a wide range of issues affecting paediatric training, and to feed back thoughts and ideas to the School Executive Team and School Board. This will be a two-way process with members of the Executive Team attending trainee committee meetings and two trainees attending the School Board. It is hoped that these experiences will help individuals within the trainee committee to develop important skills in leadership, management and medical education.

What are the different roles within the trainee committee?

The proposed make-up of the trainee committee (see Figure 1) has been built around the management structure of the School so that there is trainee input in every key area. The idea is that trainees can involve themselves in a particular area of work for a 2-3 year period, giving assistance to and learning from their senior colleagues within a particular work-stream group. The key areas requiring trainee input are detailed below:

1. Selection & recruitment
2. Workforce planning
3. Service safety / Working Time Directive / Rotas
4. Curriculum delivery
5. Assessment Strategy
6. Communication & Website & IT
7. Supporting trainees
8. Supporting sub-specialists (inc Grid / Academic / OOPE trainees)
9. Supporting audit and regional research (ie co-ordinated cross-Trust approach to audit / questionnaires etc)
10. Managing the training programmes [ST1-3, ST4-5, ST6-8]

What will be the time commitment of being part of the trainee committee?

It is planned that the trainee committee will meet every 3 months so that ideas and feedback can be taken to the Executive Team meeting that follows. There are likely to be other wide ranging opportunities to get involved with projects and ideas within your particular working area, and in addition there will be some formal leadership development training sessions for all trainee committee members. Clearly finding the time to do all of this might be problematic, but there is plenty of flexibility about when activities and meetings take place in order to minimise the disruption to clinical commitments.

It is hoped that individual trainees will stay on the committee for up to 3 years (or to CCT date if sooner) in order to have a longitudinal view of the work they are involved in.

If all of this sounds like too big a commitment but you still want to get involved with trainee representation within the School then why not consider volunteering to be the 'School Trust Rep' for your current department?

‘School Trust Reps’ – Questions & Answers:

In addition to the ideas for the Trainee Committee detailed above it is proposed that there is a trainee at every Trust who agrees to be the ‘School Trust Rep’.

What is the role of the School Trust Reps?

The value of having a trainee in each trust who is the School’s ‘Trust Rep’ has been brought up at each trainee consultation to date. The thought is that this role would ‘mirror’ that of the College Tutor, but at a trainee level. These trainees will provide a vital conduit for information to and from the School and the main body of trainees working in each individual trust.

How would this person be chosen?

The proposal is that each trust should have a ‘School Trust Rep’ who is a trainee nominated / elected / asked to participate by the College Tutor in each trust. The decision of how this person is chosen rests with the College Tutor. (Clearly it is important that this is not automatically the most senior SpR as a range of levels of trainee is clearly beneficial).

What happens when a trainee moves on to a new Trust?

Once a trainee acting as ‘School Trust Rep’ moves on to a new trust, it is the responsibility of the College Tutor to find a trainee willing to be the new ‘School Trust Rep’, to brief them about what is involved and to inform the School of this change.

How would the School Trust Reps feedback ideas to the School?

This group of ‘School Trust Reps’ will be in close (predominantly e-mail) contact with the Trainee Committee and hence the School Executive Team, thus providing a really

strong alternative route of communication in both directions. In addition ‘School Trust Reps’ will be invited (alongside all of the College Tutors) to a School Forum every 3 months. This will give them an opportunity to meet with other trainee reps and College Tutors, learn more about the major issues within the School and give ideas and input into future plans. ‘School Trust Reps’ will be expected to feedback the news and discussions from the School Forums to all of their fellow trainees at their trust.

If I am a ‘School Trust Rep’ can I also apply to be on the Trainee Committee?

It is fine for a ‘School Trust Rep’ to also be someone who has a role within the Trainee Committee, though ideally this would be someone else as it would be great to get as many different trainees involved as possible.

How do I apply to be a ‘School Trust Rep’?

There is no application form for this role as arrangements will be made locally by the College Tutor for each trust, so they are the people to talk to if you are interested in becoming the ‘School Trust Rep’ for your trust.

What will I gain by being a member of the trainee committee?

As well as being in a fantastic position to represent your colleagues and help shape the future direction of paediatric training within London, there will be wide ranging opportunities to develop your leadership and management skills. In addition to the experiential learning of working with senior colleagues on different projects the whole trainee committee will be part of a pilot study where the group is given a programme of leadership and management training. This will be based on the ‘doc-reps’ work previously run at GOSH and other research models for leadership development.

Who can apply to be part of the trainee committee?

Any paediatric trainee (ST 1-8 or SpR) who has a London training number, and those with a KSS number who are ST 4-8 or SpRs.

It is vital that we get a broad range of representation on the committee, so please do not think that this only open to 'senior' trainees. Selection will be based on application forms – the two main aims of the selection process will be to (a) get as many enthusiastic and committed trainees involved as possible and (b) ensure there is a broad representation across the region and different levels of training.

How do I apply to be part of the trainee committee?

You need to complete the attached application form and e-mail it to:

paediatrics@londondeanery.ac.uk

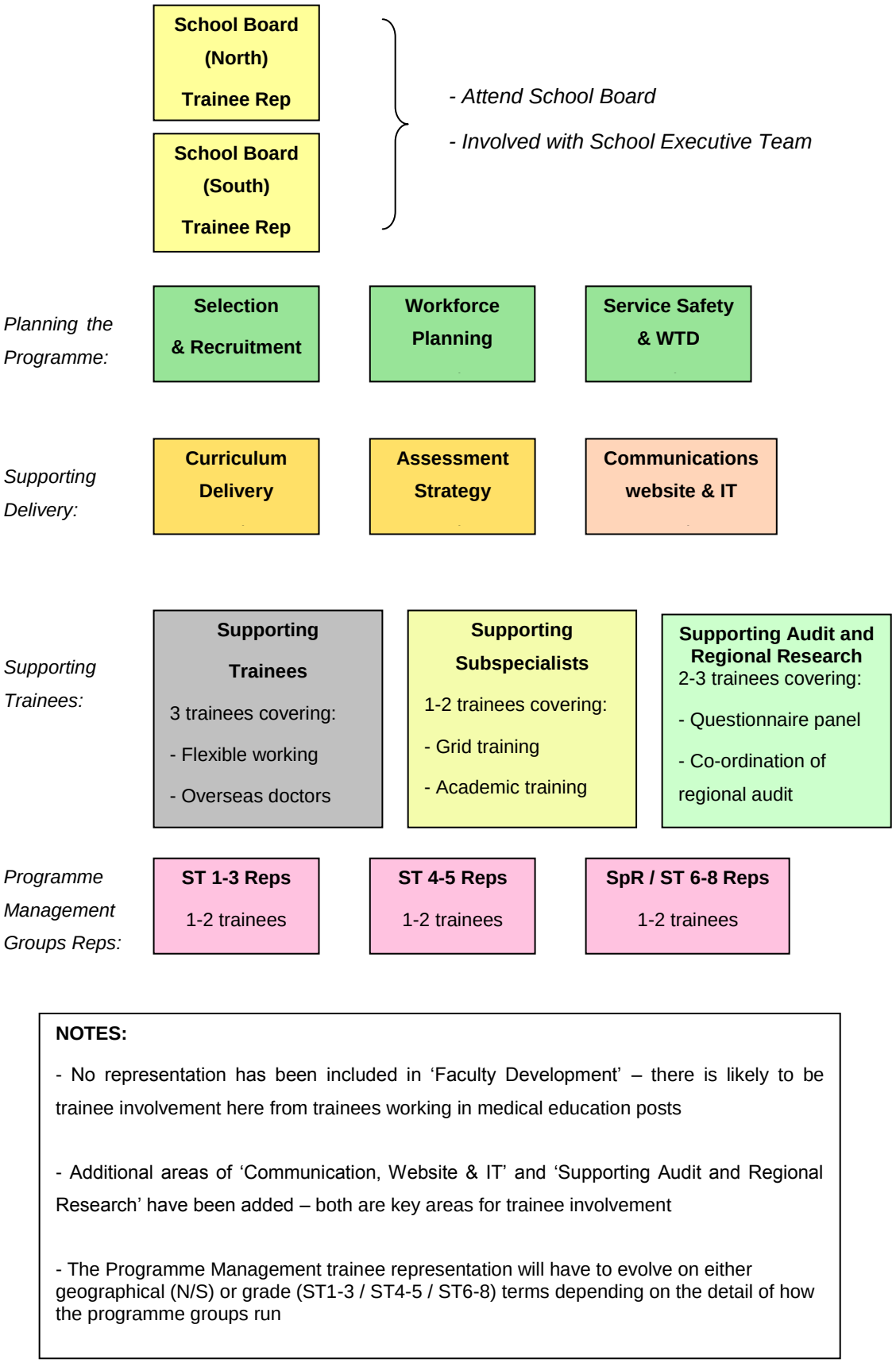
(keeping *'Opportunities for Trainee Involvement'* in the subject title)

How can I get further information about the trainee committee?

For more information you can contact:

Bob Klaber at bobklaber@doctors.org.uk

School of Paediatrics Trainee Committee – Proposed Make-up



NOTES:

- No representation has been included in 'Faculty Development' – there is likely to be trainee involvement here from trainees working in medical education posts

- Additional areas of 'Communication, Website & IT' and 'Supporting Audit and Regional Research' have been added – both are key areas for trainee involvement

- The Programme Management trainee representation will have to evolve on either geographical (N/S) or grade (ST1-3 / ST4-5 / ST6-8) terms depending on the detail of how the programme groups run