Leadership development in the UK National Health Service: Assessment of transformative academic education

Sandra Fielden

Manchester Business School

University of Manchester, UK

Claire Harris

Manchester Business School

University of Manchester, UK

Address for correspondence: Sandra Fielden, Manchester Business School, University of Manchester, Booth Street West, Manchester, M15 6PB, UK.

Tel: 0161 306 3446

Fax: 0161 306 3505

e-mail sandra.fielden@mbs.ac.uk.

Submitted to *Journal of Quality and Standards*, Revised June 2007

BSi

Abstract

To-date, unlike other training organisations, academic institutions have not been required to assess the effectiveness of the methods beyond those of academic achievement. However, this position is likely to change as academic institutions become increasingly involved in the development of bespoke programmes for both public and private sector organisations. This paper aims to address that deficit through the evaluation of an academic programme, aimed at Human Resource professionals in the National Health Service, that utlized the techniques developed for organisational training programmes in assessing its impact on learning, behaviour and results. Based on Kirkpartick's (1959) model of evaluation, the approach taken not only measures the impact of the programme on individual learning, attitudes and behaviour but attempts to evaluate how that impact translates into organisational benefits. Furthermore, it explores the on-going development of the assessment process and ways in which that process can be refined to elicit more accurate data particularly at the organisational level.

Keywords: Evaluation, Kirkpatrick, English National Health Service, Leadership Development programme.

Introduction

In 2000 the Government launched its strategy for the English NHS – The NHS Plan (Department of Health, 2000). This set ambitious targets for modernising the health service in England. Effective leadership was cited as one of the key drivers in achieving the changes needed to deliver the NHS plan "delivering the Plan's radical change programme will require first class leaders at all levels of the NHS" (Department of Health, 2000, p86). At roughly the same time, research emerging from studies in the NHS indicated the need to develop peoples transactional as well as transformational leadership competencies (Alimo-Metcalfe, 1999), namely to deal with the uncertainly of change during the modernization process. As a result the Leadership Qualities Framework (LQF) (NHS Leadership Centre, 2002) was launched, which focused on three broad clusters of competencies (setting direction; personal qualities and delivering the service).

In 2000, the human resource directorate of the Department of Health commissioned a leadership programme aimed at developing leadership skills of human resource professionals. In order to incorporate the demands of the service with academic requirements, the Leadership through Effective Human Resource Management (LTEHRM) programme was developed. The programme was adopted by the Leadership Development Centre in 2001 and widened out to other leaders (across disciplines and England, Wales and Northern Ireland) to enable them to develop their leadership skills.

Evaluation was a key element of the programme for several reasons; to assess the value to participants and the wider NHS, to ensure continuous improvements could be built into future provision, to add to the dearth of published accounts of other

similar programmes (Boaden, 2006). Kirkpatricks (1959) model was used to evaluate the programme in terms of reactions, learning, behaviours and results. This paper describes the impact of the programme at all four levels and maps the learning, behaviour and results onto the leadership qualities as outlined in the LQF.

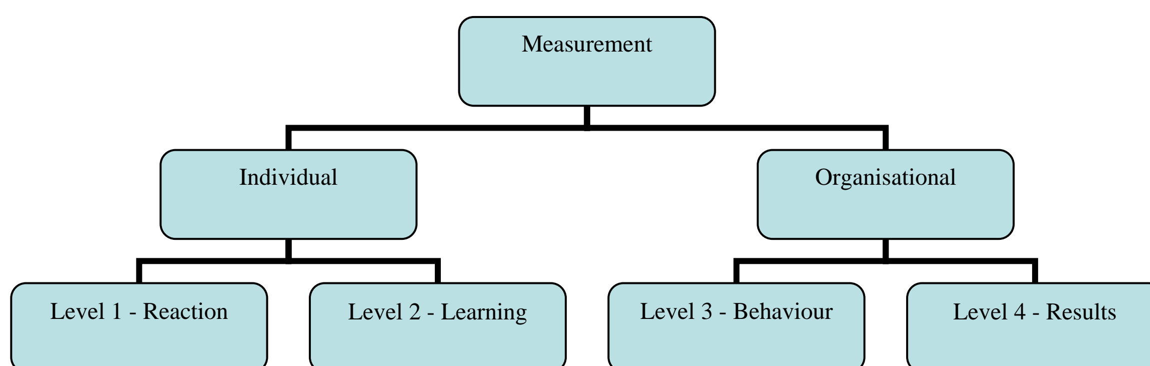
Theoretical framework

The evaluation of any learning programme is essential, not only to ensure that the client is receiving value for money but also to ensure that it is contributing to the core business, in this instance leadership development as outlined in the LQF (Swanson & Dobbs, 2006). Evaluation is the systematic collection of data in order to demonstrate the achievement of training objectives and whether that achievement has resulted in enhanced performance of the individual on the job (Collins, 2002; Phillips, 2003). Organisations are constantly challenged to devise better approaches to the evaluation of learning, so that only essential information is obtained as efficiently as possibly with the least possible cost (Stone & Watson, 2001). Forward-looking organisations view learning as a valuable investment to be monitored, with evaluation and assessment closely linked with organisational objectives, standards and competency profiles (Russ-Eft & Preskill, 2005). In a knowledge age, intellectual capital is acknowledged as an important means of acquiring a leading edge and evaluation can help support the achievement of corporate purpose by linking learning and an organisations strategic objectives or business plan (Phillips & Phillips, 2002; Swanson & Dobbs, 2006; Todesco, 1997).

Evaluation of leadership development is not a new phenomenon and the literature on evaluation has been based on a number of conceptual models and/or

frameworks, such as Kirkpartick's (1998) four-level framework, Phillip's (2003) five-level model, Swanson and Holton's (1999) assessment of performance and learning, and Baldwin and Ford's work on transfer of training (1988). However, due to the complexities inherent in the later approaches Kirkpatrick's Four Level Evaluation was adopted as the method of evaluation for this study (see figure 1).

Figure 1 – Kirkpatrick's Evaluation Framework (1998)



First developed in 1959, this model is acknowledged by many practitioners as the standard in the field, providing a vocabulary and rough taxonomy for criteria (Alliger & Janak, 1989; Naugle, Naugle & Naugle, 2000). This framework provides a solid basis for the examination of learning and development on the individual and the organisation (Watkins et al, 1998). However, it has also received a great deal of criticism for being prone to misunderstandings and over generalizations (Alliger et al, 1997; Holton, 1996), and the issues relevant to this study are discussed below in terms of each of the four levels.

Level one (*reaction*) is aimed at accessing participants' views and feeling towards the learning process, i.e. do they have positive feeling about the instructor,

the material, and the experience (Kirkpartick, 1998). The premise is that early measures of satisfaction are an indicator of programme success (Pangarkar & Kirkwood, 2006). However, the expectation that learning is associated with positive feelings of satisfaction is dangerous because courses may be modified to accommodate the desire for participants to experience greater enjoyment to the detriment of their learning. Further, this implies that learning is passive rather than active; referring to the common belief that it is the responsibility of the learning providers to ensure that participation learning occurs (Dixon, 1987). This is the most commonly used element of Kirkpartick's framework (Sugrue & Kim, 2004), yet measuring reactions rarely takes into account the participant's role as part of the learning process, asking about the course design and delivery but not assessing the participants' desire or efforts to learn.

Level two (*learning*) seeks to measure the degree of knowledge acquired, skill improvement and attitude change achieved during the course. This level of evaluation requires post-testing to ascertain skill and knowledge acquisition and retention, focusing on whether a participant can use the techniques they have acquired (Antheil & Casper, 1986; Pangarkar & Kirkwood, 2006). Kirkpatrick (1998) defines learning as "the extent to which participants change attitudes, improve knowledge, and/or increase skill as a result of attending the programme" (p.20). He notes that most learning results in some measurable gain in at least one of these three areas, with learning ideally leading to increased knowledge of concepts, the development of relevant skills and changes in attitudes towards pertinent areas, e.g. diversity. However, as all variables involved in this development process can not be identified

or their impact measured, claims of direct causality are almost impossible to make (Holton, 1996).

Level three (*behaviour*) aims to assess and evaluate whether or not participants have gained knowledge, skills and techniques, this level seeks to measure the degree to which learning is applied in the work place (Coldwasser, 2001). Kirkpatrick (1998) describes this as "a measure of the extent to which participants change their on-the-job behaviour because of the training" or programme (p21). It is important to measure behaviour, as the primary purpose of most programmes is to improve results by changing behaviour (Collins, 2002), as new learning is of little use to an organisation unless participants actually use the new skill, attitudes and knowledge in their work activities (Clark, 1997). A direct causal link is again almost impossible to make in interpreting the findings of such an evaluation and Todesco (1997) raises one particular methodological concern that demonstrates this issue. She points out that one factor that often inhibits the transference of learning is the organisational climate, where many influences are at play. Inconsistencies in that climate can undermine the effectiveness of learning programmes and it is important to identify any extraneous variables as part of the evaluation process. Furthermore, Nanda (1988) claims that it is the attitudes of top management that are the key factor in the skill development of managers, and must be taken into account when evaluating the transference of learning.

Kirkpartick's fourth level of evaluation (*results*) is a measure of the final result or impact on the organisation that occurs from the learning process. It is interdependent with level three and is concerned with direct changes in business

outcomes that are directly related to organisational strategy (Pangarkar & Kirkwood, 2006), or in this study the qualities specified in the *LFQ*. These bottom-line results are defined in terms of costs, higher quality, improved productivity, lower absenteeism and decreased turnover. Although measuring the effectiveness of programmes in these terms can be extremely useful, it is widely recognised that collecting level four information can be difficult, time-consuming and more costly than the other three levels, and there are difficulties in separating out the learning from a multitude of other variables that can have a long-term impact on performance. Numerous reviews have been undertaken of the techniques employed to capture level four evaluation (e.g. Hung & McLean, 2007; Todesco, 1997). Yet, while the econometric and utility models have statistical elegance, they fail to capture the impact of the complex products of learning: these include the impact of changing attitudes, values and ethics. However, the importance of these results should be viewed in the full context of their value to the organisation (Phillips, 2003; Russ-Eft & Preskill, 2005).

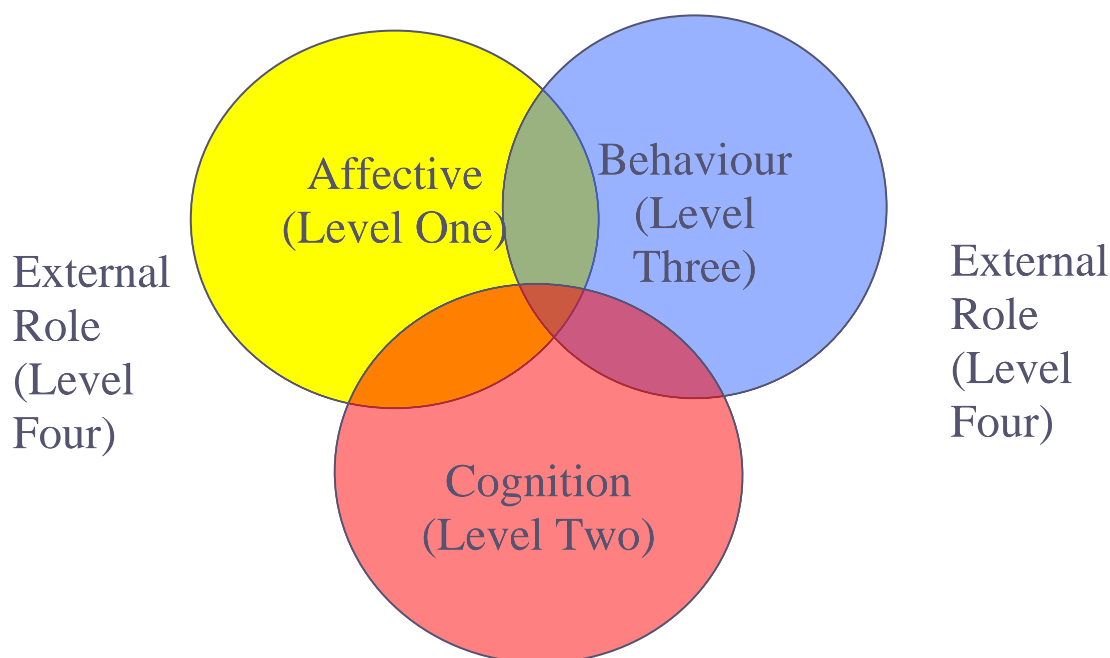
The Current Evaluation

The evaluation methodology used in this study goes far beyond the standard assessment of an academic programme (Bober & Bartlett, 2004; Torres, 2004), focusing on all levels of the Kirkpatrick model. In doing so it sought to employ a systematic approach and sought to employ varied methods in its analysis of levels three and four of Kirkpatrick's framework (1998), taking particular note of Boverie at al's (1994) suggestion that a more qualitative, quasi-experimental approach involving action research, critical incidents, and similar methods have a great deal of promise in

providing a more reliable and valid measure of the impact of learning on organisational activities.

Due to its unusual application to an academic programme the methodology itself has required testing and this paper explores the findings emerging from that embryonic approach. Its primary concern was with measuring affective, behavioural and cognitive changes experienced by participants in terms of the *LQF* as a means of evaluating the programme at the four levels of Kirkpatrick's framework (1998), while exploring how those changes have impacted on the participants own organisation and the wider NHS. The model in Figure 2 formed the basis for the approach taken, recognising the links and overlaps between each of the areas under investigation, while taking account of the influence of the external environment on learning and skill acquisition.

Figure 2 – Programme Evaluation Framework



The Leadership Qualities Framework (LFQ)

Assessment of leadership qualities and implications for the NHS highlighted that leadership development was haphazard (Alimo-Metcalfe and Lawler, 2001), yet it is widely recognised that leadership can have an impact on organisational success (Alimo-Metcalfe and Alban-Metcalfe, 2002). In 2002 The Leadership Centre published the LQF. There are 15 characteristics within the framework, clustered around personal qualities (self-belief, self-awareness, self-management, drive for improvement and personal integrity), setting direction (broad scanning, intellectual flexibility, seizing the future, political astuteness and drive for results) and delivering the service (leading change through people, holding to account, empowering others, collaborative working and effective and strategic influencing). These characteristics are reported to distinguish highly effective leaders that achieve successful outcomes across the NHS (Bolden, Wood & Gosling, 2006).

The programme

The aims of the programme in this study were to enable HR professionals and other leaders to develop their leadership skills. More specifically:

- To develop an understanding of the needs and development of the NHS
- To develop HR professionals and other NHS leaders as experts in the management and development of people
- To enable participants to develop understanding and skills in building organisational and individual capability for change
- To enable participants to develop understanding and skills in delivering change in organisations

- To enable participants to identify and address their own development needs

These aims and objectives required individual learning and behaviour change around issues such as change management, leadership and people management with the intention that they could promote lasting organisational change.

The programme was delivered by a consortium of partners, comprising Manchester Business School (MBS), PricewaterhouseCoopers (PwC) and Harvard School of Public Health at Harvard University.

The programme comprised of 3 formal requirements:

- Six “core modules” which incorporated six half week modules of teaching (at MBS) every two to four months (residential at MBS)
- Five “service improvement projects” (SIPs completed by participants after each of the modules)
- Ten “learning sets” (5 facilitated by PwC and 5 facilitated by MBS)

Participants registered onto the programme to complete a certificate in human resource leadership, with the option to proceed onto either a diploma or MSc following successful completion of the three programme requirements. Participants were also able to undertake an optional Harvard elective at Harvard University (although this did not form part of the assessed certificate).

Teaching on the “core modules” was delivered by MBS, invited speakers, employers and experts from PwC. The modules covered:

- Module 1 - Introduction to the programme
- Module 2 - Today’s healthcare agenda
- Module 3 - Leadership and change
- Module 4 - Improving the patient experience: making it happen through people
- Module 5 - Improving the patient experience: making it happen through organisations
- Module 6 - Assessment of personal and organisational development

Following each module (except for module 1) participants were required to produce a service improvement project, relevant to what they had learned from the module and applicable to their own organisation. These projects enabled participants to take their learning back to the workplace and apply it to real organisational issues. The SIPs were assessed by members of MBS.

Participants also attended two learning set days between each module, one facilitated by a member of MBS which aimed to help participants develop their ideas for their SIPs and the second facilitated by a member PwC which focused on personal development.

The participants

Approximately 400 people enrolled onto the programme (approximately 90 people were enrolled in each of the five cohorts, except for cohort 2 which had 45 participants) between 2001-2006.

Method

This paper provides results from an evaluation of the first three cohorts of the programme (approximately 250 people in total). The programme was evaluated at all four levels of Kirkpatrick's (1998) model (see table 1).

Table 1: Evidence of reaction, learning, behaviour and results from the LTEPM programme mapped onto Kirkpatrick's four levels of evaluation.

Kirkpatrick Level	Objective	Evidence
Level 1 – Reaction	Measure participants initial response to the programme	Participants rated each module on a scale of 1-5 (1= very good and 5 = very poor).
Level 2 – Learning	Measure participants learning from the programme	Participants were required to complete 5 service improvement projects (SIP) to demonstrate how knowledge gained related to their practice. Reflective SIP number 6 required participants to describe how they had related their learning from the programme back to the workplace.
Level 3 – Behaviour	Measure participants behaviour change on the job	Participants were required to write a reflective SIP (SIP 6) about their behaviour change (as a result of the programme).
Level 4 – Results	Measure participants impact on the business	Participants were required to deliver a presentation about the impact that they had made on their trust and the wider NHS.

In order to assess participants' initial response to the programme *(level 1)* they were asked to rate each taught session of the programme at the end of each of the 6 modules on a 1-5 scale (1 = Very Good to 5 = Very Poor).

Level 2 (learning) and *Level 3* (behaviour) was assessed through inspection of module 6 service improvement project (SIPs) in which participants were required to describe in detail their knowledge development (*level 2*), how this translated into behaviour change (*level 3*) and how they related their knowledge and behaviour change back to the programme. In addition, as part of the selection process all participants, their sponsor (usually their senior manager) and referee (usually their line manager) were required to rate participants against the competencies outlined in the NHS Leadership Qualities Framework (LQF): setting direction, personal qualities and delivering the service, at the beginning of the programme. These competencies were revisited at the end of the programme in order to aid participants in assessing the behavioural impact of the programme. This was also detailed in SIP 6 and so was subject to analysis.

In addition, in order to assess the impact of the programme on the organisation (*level 4 - results*), participants were required to undertake a 45 minute presentation exploring the actual or projected cost benefits that they attributed to the knowledge and behavioural changes identified in their report. The focus was on the quantifiable results of learning, in terms of tangible benefits and impacts on targets and was a measure of the final result or impact on the organisation that occurred from the learning process. This broad category was concerned with the impact of the programme on the wider community, addressing the question *"is the programme yielding value for the organisation?",* and participants were encouraged to consider the impact of the programme on their own organisation, patient care and the wider NHS. The presentations were assessed to explore what type of results had come about from the programme.

Module 6 SIPs and presentations were analysed using a pre-defined template (i.e. learning, behaviour and results) and were rated by two independent raters, thereby assuring marking consistency. For the SIPs the independent raters (academics associated with the LTEPM programme) read through each participant's module 6 SIP and highlighted evidence of learning, behaviour and results level outcomes. For the presentations independent raters (academics associated with the LTEPM programme) noted evidence of results level outcomes as the participant delivered their presentation. Any discrepancies were discussed until agreement could be met.

Results

Level one (reaction) was typically positive with over 85% of sessions rated between 1.1 and 1.9 (indicating a rating of very good – good). Less than 5% of sessions were rated as extremely poorly, i.e. those rated below 3, all of which were replaced for subsequent cohorts thereby minimising future poor ratings.

In terms of *level 2* (learning) it is important to note that approximately 30-40 participants from each cohort of participants were able to and chose to proceed onto the MSc stage of the programme (which involved writing a dissertation). This indicates that around 50% of each cohort had gained sufficient knowledge to pass their SIP assessment – demonstrating that the programme had enabled some degree of learning from the programme. However this is a conservative estimate of the number of participants that actually achieved the minimum learning requirements, as some participants chose not to continue onto MSc stage of the programme.

At the time of analysis 202 participants had completed their module 6 SIP reflecting on how the programme had impacted on their learning and behaviour however more often than not participants also reported results level outcomes, so they are also reported (*level 2* – learning, *level 3*- behaviour, *level 4* - results).

Table 2, 3 and 4 shows the percentage of those participants who reported learning, behaviour and results from the programme in their module 6 SIPs. We have also mapped the learning, behaviour or result outcome onto the LQF to show not only how being on the programme has results in all three levels of outcome but how these outcomes may also have helped to develop leadership competencies as outlined in the LQF.

Table 2: Levels 2 (*Learning*) (code refers to learning outcome 1, 2, etc; theme refers to respondents stated learning outcome; % response refers to the percentage of people that mentioned the learning outcome and LQF refers to the leadership competencies that may have been developed as a result of the learning outcome).

Code	**Theme**	**% response N = 202**	**LQF**
L1	I understand the NHS agenda better *"The key objective for me at the start of the programme was to improve my knowledge of the wider NHS agenda. This objective has been met".*	79	Seizing the future; broad scanning
L2	I have a bigger/broader/wider view *"The programme has provided me with the big picture of the NHS".*	48	Seizing the future; effective strategic influencing
L3	I have more insight into myself *"For me the benefits have been wide ranging and have included opportunities for introspection".*	20	Self awareness; self management
L4	I think out of the box *"The taught modules in conjunction with the projects have facilitated the ability to look outside the box".*	18	Seizing the future; intellectual flexibility

Table 3: Levels 3 (*Behaviour*) (code refers to behavioural outcome 1, 2, etc; theme refers to respondents stated behavioural outcome; % response refers to the percentage of people that mentioned the behavioural outcome and LQF refers to the leadership competencies that may have been developed as a result of the behavioural outcome).

Code	Theme	% response N = 202	LQF
B1	I am more confident *"The programme has helped me grow in confidence".*	100	Self belief
B2	Improved networking with HR people *"A really positive implication has been the development of a mutually supportive network of colleagues through the learning set".*	69	Political astuteness; leading change through people; collaborative working
B3	The patient is the focus of everything I do *"I have been able to use the knowledge gained to ensure that there is a clear link between the needs of the service users and HR policies and processes".*	69	Drive for improvement; intellectual flexibility; drive for results
B4	I am more motivated *"The personal and professional impact of the learning has been a quantum leap in confidence and a high degree of motivation".*	54	Self belief; drive for improvement
B5	I have been able to enable strategic development *"I have Improved my ability to develop organisational strategies that have changed the direction of the Trust."*	51	Seizing the future; broad scanning; political astuteness; drive for results
B6	I use an evidence-based approach now *"I have also learnt skills in how to influence effectively particularly using an evidence-based approach, e.g. board decisions".*	46	Broad scanning
B7	I challenge more *"Whole programme has helped build confidence to challenge the norm".*	41	Self belief; personal integrity; holding to account
B8	Leadership skills have been developed *"I think that patients have benefited from some of the benefits…in particular my enhanced leadership skills".*	38	Leading change through people
B9	I delegate more *"I have delegated work to my staff".*	8	Leading change through people; holding to account; empowering others; effective and strategic influencing; collaborative working
B10	I take more risks *"Willing to risk-take – uncharted territory".*	8	Self awareness; self belief; seizing the future;
B11	I have worked outside my comfort zone *"The learning sets challenged my learning and stretched me beyond the comfort zone".*	8	Self belief; drive for improvement; seizing the future
B12	I have become more assertive *"On a personal level I have become more assertive".*	8	Self belief; holding to account; empowering others; effective and strategic influencing
B13	I now have/am seeking a mentor *"When I return to work (after maternity leave) in whichever role, I will pay for regular quarterly coaching to help me work at this learning plan".*	8	Personal integrity
B14	I can access resources (financial) better *"Increased ability to attract funding for a variety of projects from different sources".*	8	Drive for results

Table 4: Levels 4 (*Results*) (code refers to result 1, 2, etc; theme refers to respondents stated results; % response refers to the percentage of people that mentioned the result and LQF refers to the leadership competencies that may have been developed by achieving the stated result).

Code	Theme	% response N = 202	LQF
R1	HR as a function rated more highly/higher profile within the organisation *"The programme has raised the author's credibility and profile in the organisation and has highlighted that HR can and should take the lead in effecting change".*	44	Political astuteness; leading change through people
R2	I have changed my Job/Role *"The programme has been beneficial in terms of personal development; career development as the author has taken up a Director post sine the programme began".*	38	Self belief; drive for improvement
R3	I have increased credibility with clinicians *"While I have always has reasonable credibility amongst my colleagues and peers I now experience greater involvement and respect from clinical academics than I have before".*	31	Political astuteness; leading change through people; holding to account; empowering others; effective and strategic influencing; collaborative working
R4	I have been able to share best practice *"There are many examples of sharing good ideas within the learning set, e.g. on flexible working, bullying and strategy development".*	23	Intellectual flexibility; broad scanning; effective and strategic influencing
R5	The programme has retained me within the NHS *"Most importantly, and I could not have predicted this at the start of the programme, it has probably been the main contributor to my retention".*	18	Drive for improvement
R6	The investment has had a high payoff *"I am able to demonstrate significant cost benefit as a result of my attendance on the programme".*	18	Drive for results
R7	The HR community developed as a community *"These networks are invaluable ensuring cross fertilisation of ideas and best practice across the NHS as well as developing a real sense of an HR community that looks outwards rather than within".*	8	Collaborative working

At *level two* (learning) over 70% of participants reported that they had gained a better understanding of the NHS as a result of being on the programme. At *level three* (behaviour) 100% of participants reported an increase in confidence, and over 50% reported that they were more motivated, were better able to develop strategy and network with other people and that they had adopted a patient focus in the practice. This is in line with the NHS strategy to provide a patient focused NHS. Around a third of participants reported changing job roles as a result of the programme and 18% reported that they had remained in the NHS as a result of being on the programme.

Level four (Results) was assessed by recording examples of cost benefits provided by participants during a 45 minute assessed presentation, in total 207 participants undertook the presentation. Although some tangible examples of cost benefits could be directly attributed to the presenters' participation on the programme, the majority of examples could only be indirectly attributed to the programme and participants were reluctant to provide any cost analysis of these examples. This was mainly because presenters felt that there were too many extraneous variables for a direct linkage to the programme to be supported. However, there was a general recognition that it was the development of participants at the individual level which was responsible for at least some of these cost gains.

Examples of *level four* achievements included: decreased staff turnover of between 1%-2%, securing an external grant of £500 000 for day care facilities, and cessation of exit interviews saving approximately 900 staff hours. Actual cost benefit analysis was extremely restricted as the majority of participants (80%+) felt that it was too early to give an accurate assessment of level 4 attainments, i.e. only 2 months after the final taught element of the programme.

In terms of the LQF, all of the reported outcomes from the programme can be mapped into the LQF competencies (see table 5).

Table 5: Evaluation outcomes of LTEMP programme mapped onto the LQF.

LQF competence	Evaluation outcome	Number of times mapped
Setting direction		
Seizing the future	L1, L2, L4, B5, B10, B11	6
Intellectual flexibility	L4, B3, R4	3
Broad scanning	L1, B5, B6, R4	4
Political astuteness	B2, B5, R1, R3	4
Drive for results	B3, B5, B14, R6	4
Total		**21**
Personal qualities		
Self belief	B1, B4, B7, B10, B11, B12, R2	7
Self awareness	L3, B10	2
Self management	L3	1
Drive for improvement	B3, B4, B11, R2, R5	5
Personal integrity	B7, B13	2
Total		**17**
Delivering the service		
Leading change through people	B2, B5, B8, B9, R1, R3	6
Holding to account	B7, B9, B12, R3	4
Empowering others	B9, B12, R3	3
Effective & strategic influencing	L2, B6, B9, B12, R3, R4	6
Collaborative working	B2, B9, R3, R7	4
Total		**23**

Discussion

The evaluation process has demonstrated outcomes associated with the LTEPM programme are at every level of Kirkpatrick's framework, although no causal link can be proven. Participants' initial reactions to the programme were favourable, and the number of people completing the formal elements of the programme (i.e. taught, learning sets and service improvement projects) and progressing onto higher academic qualifications indicates a good level of learning has been achieved. Further, analysis of the module 6 SIPs and presentations shows that people were able to provide evidence that what they had learned, changed their behaviour and produced organisational level results due to their participation on the programme, reporting measurable changes in terms of personal, career and organisational impact. It is clear to see that participants reported increases at all levels of development (i.e. learning, behavioural and results), although overall they were more comfortable reporting tangible organisational impacts than more reflective personal ones. This may not be surprising as only 54% reported that they had gained learning in the form of greater personal insight. However, the benefits to the organisation are highlighted with over half of participants beginning to practice patient focused HR, taking a more strategic approach to HR and improving their networking with other HR professionals.

In terms of the LQF '*Delivering the Service*' (mapped 23 times) was reported as the area where participants experienced most development, closely followed by '*SettingDirection'* (mapped 21 times) and the *'Personal Qualities'* (mapped 17 times). As Swanson & Dobbs (2006) rightly points out, the success of an evaluation is to demonstration of the programme has delivered its objectives and contributed to the

core business, in this case leadership development. These findings clearly show that the programme has achieved the desired outcomes.

However, this form of evaluation can only demonstrate change by those who completed all three elements of the programme. It could be argued that participants were self-selecting in this regard, as the choice as which elements of the programme to engage in was ultimately theirs. Thus, while the data provides a good picture of those individuals who were motivated to learn or who had a high need for achievement, it does not provide any evidence around the development of those who did not take part in the evaluation process above *level one* (reaction).

The main criticism aimed at level two as an evaluation method is that is does not demonstrate how learning is applied on the job (Antheil & Casper, 1986). Level three (*behaviour*) sought to address this omission by exploring the transfer of learning into the workplace. This form of evaluation requires both pre and post-testing to measure the degree of knowledge acquired, skill improvement and attitude change. In order for the measurement tool to be most effective it needed to reflect the programme's objectives (Collins, 2002; Phillips, 2003) and competencies are recognised as one of the most effective and informative measures of learning. The competency data collected during the participant selection stage of the programme was ideal, in that it provides quantitative data that relates directly to the programme's objectives. In theory, competency comparison would have allowed for statistical analysis to be conducted which, according to Kirkpartick (1996), is essential if evaluation is to be accurate and meaningful. However, the degree of career movement experienced by participants during the course of the programme was unexpected. This arose for two reasons, firstly because of the direct impact of the

programme on the career paths of participants and secondly because of the extensive reorganisation of the NHS that took place during the programme. The result of this movement meant that the majority of those participants who took part in *level three* (behaviour) evaluation, could not obtain post-competency ratings from their original sponsor and referee. Thus, competency shifts could not be statistically analysed and could only be used as an indication of change.

Evaluation at *level four* (results), as with the suggested level five assessment (Phillips, 2003), is extremely time consuming and it is difficult to ensure that the findings are meaningful. Evaluation of the cost benefits gained from the LTEHRM programme was even more difficult because of the type of learning involved. Some participants were able to provide examples of behaviour change that had resulted in increase income, e.g. through government grants, or decreased expenditure, e.g. through reduced staff turnover, for their organisation. However, for the majority of participants the career moves they had experienced meant that the isolation of actual cost benefits was not possible at the time of assessment. Thus, it is suggested that level four outcomes (*results)* should be reassessed after 12-18 months, in order to more fully assess the bottom-line impact of the programme. Evaluation at this level must also take into account economic, political and demographic variables that influence organisational performance, in order to accurately assess change (Todesco, 1997).

Although analysing the participant's personal and organisational development reports was a useful exercise, in the future it is recommended that other techniques for evaluating the LTEHRM programme at each of the Kirkpatrick's four levels framework be adopted (Boverie et al, 1994). It is recommended that in addition, data

are collected relating more directly to factors influencing motivation and participation, providing a more in-depth understanding of how these factors impact on the programmes effectiveness to achieve change at all levels. These would focus not just on the individual participant but how their experiences, enjoyment and involvement were influenced by their peers, work colleagues, family members, work demands. Data collection would include investigation into: action leaning set experiences (e.g. benefits, impact of peer expectations), personal stressors or inhibiting factors (e.g. care responsibilities), career changes (e.g. promotion), work stressors and job demands (e.g. additional work pressure as a result of staff shortages). In order to gain a fuller understanding of the reasons behind participant withdrawal from the programme, this additional data would need to be collected from those who fail to complete the course. Thereby providing a more accurate picture of the type of environment that is most likely to produce high performance and a subsequent transfer of learning into the workplace (Bladwin & Ford, 1988).

5. Conclusion

The approach take to the evaluation of this academic programme is far more in line with corporate evaluation than the traditional forms of academic assessment (Antheil & Casper, 1986; Bober & Bartlett, 2004). Although the methodology itself is subject to uncontrollable external influences, such as major organisational change, the developmental stages, the results reveal that this approach can and does demonstrate the positive impact of such a programme on learning, behaviour and attitudes. The relationship between academic providers and their potential client groups are changing, becoming more like those between client organisations and training

providers. Academic clients (especially those purchasing postgraduate programmes) are becoming increasingly aware that they need to justify to their stakeholders that such programmes are value for money (Bober & Bartlett, 2004). Thus, comprehensive evaluation methods need to be developed for academic programmes that can meet the growing need for effective and reliable assessment. Finally, participants themselves need to be motivated and committed to learning and development if they are to gain any real benefit from such programmes, i.e. not just attending because they have been nominated. Recognising the needs and abilities of the learners is essential, as is their motivation and commitment to the programme at all stages and aspects of learning if an accurate picture of the potential of academic programmes to impact on leadership in the workplace is to be established.

REFERENCES

Alimo-Metcalfe, B. (1999). Leadership in the NHS. In A.L. Mark & S. Dopson (eds.) *Organisational Behaviour in Health Care: The Research Agenda.* London: Macmillan Press Ltd , pp135-153.

Alimo-Metcalfe, B., & Alban-Metcalfe, R.J. (2002). Leadership. In P. Warr (Ed). *Psychology at Work.* London: Penguin.

Alimo-Metcalfe, B. & Lawler, J. (2001). Leadership development in UK companies at the beginning of the 21st century- lessons for the NHS? *Journal of Management in Medicine,* 15, 387-404.

Alliger, G. M. & Janak, E.A. (1989). Kirkpatrick's levels of training criteria: thirty years later. *Personnel Psychology.* 42 (2), 331-342.

Alliger, G.M.; Tannenbaum, S.I., Bennett, W.Jr.; Traver, H. & Shortland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology.* 50 (2), 341-358.

Antheil, J. H. & Casper, I. G. (1986). Comprehensive evaluation model: A tool for the evaluation of non traditional educational programs. *Innovative Higher Education*, 11(1), 55-64.

Baldwin, T.T. & Ford, J.K. (1988). Transfer of training: a review and directions for future research. *Personnel Psychology.* 41 (1), 63-105.

Boaden, R. (2006). Leadership development: does it make a difference. *Leadership Development,* 27, 5-27.

Bober, C.F. & Bartlett, K.R. (2004). The utilization of training programe evaluation in corporate universities. *Human Resource Development Quarterly.* 15 (4), 363-383.

Bolden, R., Wood, M., & Gosling, J. (2006). NHS leadership qualities framework: Missing the wood for the trees. In A. Casebeer, A Harrison, & A.L. Mark (eds). *Innovations in Health Care: A Reality Check.* New York: Palgrave Macmillan, pp, 17-29.

Boverie, P.; Mulcahy, D. S. & Zondlo, J. A. (1994) Evaluating the effectiveness of training programmes. In J. W. Pfeiffer (ed) *The 1994 Annual: Developing Human Resources.* San Diego, C.A.: Pfeiffer & Company.

Coldwasser, D. (2001). Beyond ROI. *Training.* 38 (1), 82-90.

Collins, D.B. (2002). Performance-level evaluation methods used in management development studies 1986 to 2000. *Human Resource Development Review.* 1 (1), 91-110.

Clark, D. (1997). Instructional system development - evaluation phase. Http://www.donclark/hrd/sat6.html.

Department of Health. (2000). *The NHS Plan: A Plan for Investment, A Plan for Reform*. London: Department of Health.

Dixon, N. M. (1987). Meeting training's goals without reaction forms. *Personal Journal*, 66(8), 108-115.

Holton, E. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7(1), 5-21.

Hung, Y. & McLean, G.N. (2007). Training evaluation based on cases of Tiawanese benchmarked high-tech companies. *International Journal of Training and Development.* 11 (1), 35-48.

Kirkpatrick, D. (1959). Techniques for Evaluating Training Programmes – part 2: Learning. *Journal of the American Society of Training Directors.* (December), 21-26.

Kirkpatrick, D. (1998). *Evaluating Training Programmes*. San Francisco, CA: Berrett-Koehler publishers, Inc.

Naugle, K. Naugle, L. B. & Naugle, R. J. (2000). Kirkpatrick's evaluation model as a means of evaluating performance. *Education*, 121(1), 135-145.

NHS Leadership Centre. (2002). *NHS Leadership Qualities Framework*. London: NHS Leadership Centre.

Nanda, R. (1988) Organizational performance and supervisory skill. *Management Solutions.* 33(6), 22-28.

Panganar, A. & Kirkwood, T. (2006). Measuring return on leadership training. *Canadian HR Reporter.* 19 (6), 14.

Phillips, J. J. (2003). *Return on Investment in Training and Performance Improvement Programmes.* San Fransisco, C.A.: Butterworth-Heinemann.

Phillips, J. J. & Phillips, P.P. (2002). Technology's return on investment. *Advances in Developing Human Resources.* 4 (4), 512-532.

Russ-Eft, D. & Preskill, H. (2005). In search of the Holy Grail: return on investment evaluation in human resource development. *Advances in Developing Human Resources.* 7 (1), 71-85.

Stone, J. & Watson, V. (2001). Evaluation of Training. http://isps-atlanta.org/Evaluation.htm

Sugrue, B. & Kim, K.H. (2004). *State of the Industry Report.* Alexandria, VA: ASTD.

Swanson, R.A. & Dobbs, R.L. (2006). The future of systemic and systematic training. *Advances in Developing Human Resources.* 8 (4), 548-554.

Swanson, R.A. & Holton, E.F. (1999). *Results: How to Assess Performance, Learning and Perceptions in Organizations.* San Fransisco, C.A.: Berrett-Koehler.

Todesco, A. (1997). *From training evaluation to outcome assessment: what trends and best practice tells us.* Learning Services Directorate Public Services Commission, Sept.

Torres, R.T. (2004). Invited reaction: The utilization of training programe evaluation in corporate universities. *Human Resource Development Quarterly.* 15 (4), 385-388.

Watkins, R. Leigh, D. Foshay, R. & Kaufman, R. (1998). Kirkpatrick plus: evaluation and continuous improvement with a community focus. ETR&D-Educational *Technology Research and Development*, 46(4), 90-96.