Northumbria Research Link

www.northumbria.ac.uk/nrl

northumbria
UNIVERSITY NEWCASTLE

Evaluating Training and Development in UK Universities: Staff Perceptions

Introduction

This paper explores research into individual experiences of training evaluation in UK Higher Education Institutions (HEIs) and identifies areas for improvement of evaluation. The research is drawn from a wider two-year project funded by the Higher Education Funding Council, under the Leadership and Governance stream which examined training evaluation in UK Higher Education to develop a diverse range of training evaluation tools for HEIs. Following Sambrook (2001), the research presented is of an applied nature, and focuses upon the perceptions of University staff experiences of how training and development is evaluated within their University.

According to UK Commission for Employment and Skills (UKCES, 2011), employers spend over £49 billion a year on learning and development interventions. It is therefore crucial that organisations can ascertain the value they gain from their investment in training (Pineda, 2010), particularly in the current economic climate. Mann (1996:14) shares a widely accepted view about the importance of evaluating learning and development interventions, "with the huge investment in developing training strategies, the question is no longer "should we train" but rather "is the training worthwhile and effective?". This is consistent with a wide range of literature which accepts that training in organisations is important and recognises evaluation as key in proving its 'worth.'

The significance of evaluation is demonstrated within existing literature, where for example "evaluation is an integral part of training" (Bimpitsos and Petridou, 2012:990). Effective evaluation has been an on-going challenge for organisations and HRD professionals (Griffin, 2011), with Lewis and Thornhill (1994:25) confirming *"there seems to be widespread agreement with the proposition that evaluation is the least well conducted aspect of all training activities."* This proposes a need for organisations to focus on evaluation of the training and development process. This standpoint was confirmed in the CIPD (2012) 'Learning and Development' survey whereby one in seven respondents revealed they did not routinely evaluate their training provision.

This current qualitative research study adds further to the knowledge base on training evaluation and addresses an on-going need within the recent evaluation research by Diamantidis and Chatzoglou (2012), who noted that a limitation of their own study was its quantitative approach and the geographical (Greek) context in which it was situated. The guiding research question addressed in our research is, what are staff perceptions of how their training and development is evaluated in two UK HEIs? The answer to this question is important as the existing research base is very clear that learners have to be at the centre of the evaluation process in order for it to be meaningful and that it has to be a *'co-creation process'* (Lingham et al., 2006:335); we address this gap explicitly within our research.

We begin with an insight into the key literature on training evaluation, highlighting the central debates. As part of our contribution we offer insights into two UK university case organisations. We then briefly outline key contextual factors and highlight the lack of specific studies in this area. A discussion of the research methods and ethical issues then follows and leads to presentation of the qualitative research findings and discussion drawing upon the extant literature. Due to our emphasis on applied research, we outline how the research findings influenced the development of a practical training evaluation toolkit for implementation by HRD professionals in UK HEIs. Thus also responding to, Griffin's (2011) call that there is insufficient research aimed at practitioners and how they can improve their evaluation practice.

Literature review

Training evaluation

Trainers aim to design and deliver training and development activities in such a way that practical benefits in the workplace can be observed and to enable the employees to transfer new knowledge and skills for the benefit of individuals/departments and the overall organisation. Training programmes are effective "only to the extent that the skills and behaviours learned and practiced during instruction are actually transferred to the workplace" (Chiaburu and Lindsay, 2008:199). There are numerou*s* understandings of training evaluation. For example, Pineda (2010:674) notes, "evaluation involves collecting information on the results obtained in order to analyse and evaluate them and facilitate the optimisation of training in the future" while Dawson (1995:3) also provides a useful but straightforward explanation,

> The evaluation of training is the systematic and impartial collection of data for managers and all other interested parties. This information equips them to draw conclusions about the effectiveness of particular training measures as a way of achieving organizational objectives, implementing policy and promoting organizational learning.

Organisational benefits of evaluation are generally agreed as: supporting the decision-making process on what interventions should be re-commissioned; providing evidence of investment in human capital and demonstrating the value that training interventions bring (Campbell, 1998; Griffin, 2011). In tougher economic times evaluation data may also be used to justify the expenditure of training and development Departments. APSC (2005:2) suggest evaluation is particularly important in the public sector, noting "effective evaluation is part of an accountable, professional and ethical public service. It is fundamental to good governance, good practice and good management". This suggests that public sector organisations should pay particular attention to evaluation to demonstrate the value and benefits accrued from their investment.

Models of evaluation

In the 1960s Donald Kirkpatrick produced a series of articles concerning evaluation where he identified four stages/ levels of evaluation (see Table 1). Despite its age, Kirkpatrick's model (1977:9) continues to be used in contemporary research (see for example, Gubbins, *et al.,* 2012).

<Insert Table One here>

The model's durability is its simplicity, designed to allow it to be understood and used easily by HRD practitioners when designing evaluation tools. There are however limitations associated with the four stages. For example, Kirkpatrick (1977) confirms that the usefulness of stage one evaluations is dependent upon the honesty of the learners who complete the questionnaire. There is also the potential issue of social desirability especially if not completed in private (Darby, 2006) and a risk that the evaluation focuses too heavily on enjoyment, reflected in the use of 'happy sheets' (Bramley and Kitson, 1994), rather than on how useful the training has been (Hamblin, 1974). Stage two evaluations provide learners with the opportunity to evaluate their training and to identify how they believe their behaviours might change. In relation to stage two, Kirkpatrick (1977) believes that there is a distinction between gaining 'evidence' and 'proof,' where proof can only be gained by comparing pre- and post- intervention performance to identify whether there has been a shift. However an outstanding limitation would be how practitioners are able to control for other factors.

Stage three of Kirkpatrick's (1977) model, 'behaviour' is understood as the extent to which behaviour has changed as a consequence of the intervention. Similar issues arise for stage three, where a control group would be needed to ascertain whether the intervention is the differing factor. Level four is understood as "the final results that occurred because the participants attended the programme" (Kirkpatrick, 1998:60) which can be more difficult to meet as it may be unsuitable for all training programmes. Inherent challenges in evaluation at levels 3 and 4 of Kirkpatrick's (1977) model are understood to be down to the other factors which may impact on changes in performance, including personal, instructional and organisational factors (James and Roffe, 2000). Alliger and Janak (1989) argue that some interventions will not aim to meet all four levels but the model can still be applied, assuming it meets the required needs of the organisation. Therefore in considering what forms of evaluation are required, organisations may identify that for some programmes, evaluation at stages one and two may be sufficient.

Dyer (1994) suggests a revised version of Kirkpatrick's (1977) model, referred to as 'Kirkpatrick's Mirror Evaluation System' which begins by looking at the final stage of the model rather than the first. Whilst an advantage of this model is that it includes the all-important link to business objectives, there is little empirical evidence of this system being in use. Phillips (2002) suggested an addition of a fifth stage to Kirkpatrick's evaluation model, 'return on investment' (ROI) to measure the monetary costs of providing the intervention with the monetary benefits received but recommended that due to the costs involved it should be used sparingly.

The foundations of evaluating training and development

An effective system of evaluation begins with a comprehensive training needs analysis (TNA) process so that the organisation, departments and individuals can identify their specific needs (Alvarez *et al.,* 2004). This TNA process can take a variety of forms: at an individual level it may be discussed as part of the performance management process; it may be a discrete activity in its own right or may be driven from other organisational data. Throughout the process attention should focus on the linkage with organisational strategies and goals to ensure that there is consistency and that training interventions have the potential to have an impact (Anderson, 2009). Clearly there is a close association with training evaluation, as programme/intervention objectives and return on expectations and/or investment should be considered as part of the TNA process and vice versa (Tennant *et al.,* 2002). The literature is also explicit about the need for all stakeholders to be clear about the purpose of evaluation and the learning objectives, otherwise they are unlikely to be addressed adequately (Campbell, 1998). Pineda (2010:685) allocates the role of providing this clarity to trainers and comments, "trainers and training specialists design the evaluation system and drive and oversee its implementation. For this reason they should obtain the cooperation of other agents and negotiate their level of involvement in the evaluation of transfer".

Designing and collecting appropriate evaluation data

The most frequently used method of data collection is requesting participants to complete evaluation sheets at the end of an event. The advantage of this method is a quick process, often based on a tick-box approach; which allows the full dataset to be compiled and summarised easily. The speed in which these forms can be completed at the end of the session is also useful for trainers as they tend to achieve high response rates. Darby (2007) established that where open-ended questions are included, there was a negative effect on response rates: a disappointing finding, as they offer richer data. Further, a common limitation with self-completion questionnaires is the tendency for the same forms to be used for every intervention, regardless of diverse content and

structure (Dawson, 1995). As well as limiting the range of information provided there is also a danger that learners get bored in completing the same form after every session and resort to 'default' responses. To overcome this, evaluation mechanisms should be tailored for each event and should be considered at the design stage of new interventions (Diamantidis and Chatzoglou, 2012). In practice this could provide a resource challenge for organisations who may decide to prioritise specific programmes where evaluation data is more critical.

Russ –Eft and Preskill (2008) suggest using alternative distribution methods to collect the evaluation data (e.g., electronically) and to limit the amount participants are expected to complete. This requires organisations to agree priorities and identify efficient ways in which information should be collected. Methods can include assessments, participant and supervisor surveys, interviews, observations and review of performance data. James and Roffe (2000) argue that some organisations can be cynical about using different mediums of evaluation, so there is a need to agree expectations and where appropriate, for the trainer to provide a business case. Practical issues also need to be considered in order that levels of reflection (and the resources required) are commensurate with those of the interventions taking place, as well as issues of learners' anonymity and confidentiality. However the bottom-line is that the "usefulness of evaluation data depends on its accessibility and usefulness to those who can effectively collect, analyse and employ it" (Weissner *et al.,* 2008:380) which highlights the need for expertise in analysing the data and commitment to take relevant actions.

There is debate between researchers in relation to the need to undertake an additional evaluation after a given period of time (e.g., Chappell, 2008) and a time lapse making it harder for learners to recall meaningful information (Short, 2009), where organisations make the best possible use of the 'window of opportunity' so that specific benefits can be isolated and explored. Evaluative 'testing' just after the event is advocated by Mann (1996) but other researchers are split over the ideal time-period for evaluation (e.g., Campbell, 1998).

Trainers play a critical role in the successful evaluation of events, yet to date there is a dearth of literature exploring their role and the impact of it. Campbell (1998) contends that they are responsible for undertaking post-event action plans, however there is little empirical data demonstrating the extent to which this takes place or how they are accountable for their actions. Pineda (2010:690) takes a different stance, noting the lack of preparation of training professionals, who feel ill-equipped to deal with the complexity that training evaluation entails.

The role of line managers in the evaluation process

A key theme within the existing literature is the role of line managers in supporting the training and development of employees; before, during, immediately afterwards and finally in evaluating the impact. Line managers are increasingly expected to take the role as developers of their staff. Indeed; "the capability and commitment of managers and key workers throughout an organisation's hierarchy are the most important factors in determining the provision and effectiveness of workplace learning" (NSTF, 2000:37). Also 40% of respondents to CIPD's (2012) 'Learning and Development' survey believed that more responsibility for learning and development would continue to be passed to line managers. This is despite wider literature on the devolvement of people management responsibilities to managers who may not have the relevant skills and knowledge (Author A and Author B, 2011). "Managers will not champion training and development if they are fearful about the achievements of their subordinates and uncertain about their own capabilities" and furthermore, their motivation towards the development of their employees will be partially based on self-interest and their own personal values (McGuire *et al.,* 2008:346).

With regards to evaluation, 43% of respondents to the Industrial Society's (2001) survey said that lack of line manager support was a core problem. Potential ways to engage line managers in training evaluation processes is to highlight the benefits that can be gained from information collected as a result. In addition to improving the learning process and decision-making, benefits include: improving relevant business areas; organisational culture; engagement with stakeholders; meeting internal and external reporting requirements and managing risk (APSC, 2005):

The UK Higher Education Sector: Staff Training and Development

The UK HE sector is experiencing unprecedented change, attributable to a range of factors, many of which are externally generated. Examples include Government-led changes to student funding arrangements and the implications of the recession, as well as the challenge of conflicting pressures for HEIs to collaborate and compete with each other (Barnett, 2003). There are also increasing legislative requirements (Gordon and Whitchurch, 2007) and high levels of external accountability (and media focus), including the National Student Survey and Quality Assurance Agency (QAA) audits which impact upon HEI staff activities and workloads.

Such challenges, changes in HEI staff work and increasing workloads require appropriate training and development support to adjust to new roles and responsibilities (Blackmore and Blackwell, 2004). Whitchurch's (2008) research on the changing role of non-academics also has implications for the design, delivery and evaluation of training interventions at a team and departmental level as she explores the increasingly diverse range of activities to be carried out and how[University staff] "are moving laterally across functional and organisational boundaries" (Whitchurch, 2008:1). Whitchurch (2008) contends that the situation in HEIs is more dynamic and multifaceted than individual's job descriptions or organisational structure charts suggest. This potentially has significant implications for training and development strategies.

At the end of July 2011 there were 381,790 staff employed in UK HEIs, of whom 118,115 (47.46%) were academics and 200,605 (52.54%) non-academics (HESA, 2012), demonstrating the size of the sector and the number of employees likely to be exposed to training and development interventions within their institutions. HEIs invest considerably in staff development and therefore

evaluation of these activities is critical to evidence progress on individual and organisational performance and development.

Having conducted a comprehensive web-based search for existing research on evaluation of training in UK HEIs, only two sector-specific reports were available. Firstly, "Evaluation of staff learning and development" at Anglia Ruskin University Library (Cefai, 2009) which focuses solely on support staff however some of their recommendations may be generalisable more widely as they focus on more general best practice including the need for pre-training evaluation (TNA) (Tennant *et al.*, 2002), post-training evaluation when learners have had the opportunity to use their knowledge and skills (Chappell, 2008) and the key supporting role of line managers (Pineda, 2010) As Cefai (2009) did not undertake primary research with key stakeholders, our paper addresses this gap. Secondly, an important contemporary study is the, "The Baseline study of Leadership Development in HE" by Burgoyne, Mackness and Williams. (2009:5) as one element of this report included a discussion of evaluation in the context of leadership development interventions, and highlighted evaluation as a common weakness, in that much evaluation is of an informal kind and "explicit and systematic evaluation is not the norm" (Burgoyne *et al.,* 2009:5). Our paper responds to this by focusing specifically on formal evaluation processes and how they are utilised (or not) within HE.

To summarise our discussions thus far, we have identified three areas critical to the effective evaluation of training interventions; have outlined the importance of it in the changing context of UK HE and have established that the evaluation process is under researched. It is against this background that we sought to explore staff perceptions in UK HEIs of how their training and development is evaluated.

Research Design

From the evaluation literature discussed earlier, it is clear that a 'one size fits all' approach to training evaluation would be inappropriate (Anderson, 2009) and that the characteristics of the UK HE sector should be taken into account when implementing evaluation. Two post-92 universities (former Polytechnics, awarded University status after 1992) committed to engage in the research and granted access to their staff. In order to address the overall research question which guides this paper, 'what are staff perceptions of how their training and development is evaluated in two UK HEIs?,*'* we designed and implemented four focus groups with staff from each university over a three month period in 2011. Focus groups are an appropriate research choice as "[they] have, over time, evolved into a major research tool applicable in all fields where behavioural influence and its impact or outcome can be measured and subsequently analysed" (McLelland, 1994: 29).

52 staff were randomly selected by the two HR departments to take part in the focus groups from the target job role groups of academic and non-academic staff, with a second criterion of having attended a formal training intervention within the previous 12 months. 47/52 participants accepted the invitation 23 from University X and 24 from University Y as shown at Table 2. The purpose of the focus groups was to identify key themes in order to inform the design and production of a toolkit of evaluation resources, rather than to examine the individual voices of participants.

<Insert Table Two here>

Protocols of organisational informed consent, anonymity and confidentiality were followed; participants provided written informed consent and were granted anonymity to ensure identities could not be identified directly or indirectly. The focus groups were facilitated by the first author

who received permission to digitally record the meetings. A note-taker was also in attendance to back-up the transcription process. The same key questions were asked in all focus groups (see Table 3).

\<Insert Table Three here\>

The resulting qualitative data was analysed using template analysis (King 2004; 2012), one of the most well-used forms of thematic analysis in the management field which can be applied to any kind of data (Cassell, 2012), to identify and compare key themes from across both universities. We followed the conventions outlined by King (2004; 2012) and were also guided by McGivern's (2009) approach which emphasises the need to begin with a *mechanical* exploration before moving on to the *intellectual analysis*. This consisted of the first author conducting an initial exploration of the focus group data to get a 'feel' for it and to devise inductively the initial thematic categories. This was followed by the 'intellectual analysis' involving both authors which involved an iterative process of moving between the data and the literature and resulted in further thematic categories. This process of coding and interpretations enables researchers to discover links within and across the categories of the template and enables conceptual or theoretical links to be made (Cassell, 2012). Examples of the data analysis are shown at Table 4.

Summary of findings

Clarity in the reason for evaluating training and development

Across all eight focus groups the participants did not demonstrate a detailed understanding of the purpose of evaluation, other than as a means of the University judging whether or not the sessions are well delivered. This is contrary to the good practice discussed by Campbell (1998) who highlighted the need for all stakeholders to have an awareness of the purpose and objectives. There was an overwhelming assumption from the participants that trainers are required to provide feedback after every session, in line with good practice, (Towler and Dipboye, 2009). However interviews with University HR representatives demonstrated that this was not a mandatory requirement in either university.

Only a small minority of participants identified that they were asked to consider learning objectives *before* they attended development sessions, and suggested that this was line-manager driven rather than a University requirement. Other members of the focus groups agreed that the use of personal objectives could add value and might enable them to see how the training could support them in their current or future job role. Caution was urged by some of the non-academic attendees who felt that a detailed TNA process might discourage attendance.

Increased engagement with personal objectives may improve the transfer of learning into the workplace so essential for the individual and the organisation (Chiaburu and Lindsay, 2008). It was interesting to note the perspective of Participant D who expressed a strong belief that individual learners need to take ownership of transferring *their* learning into *their* workplace *"It is about looking for the opportunities, it is about personal responsibility"*. Whilst their focus group participants showed agreement via body language, these comments were not verbally confirmed.

Participants from Focus Group Two agreed that attendance at training events was aligned closely to their performance development plans, which is in line with good practice (CIPD, 2012). However experiences reported from the other groups was more diverse. Participants across all focus groups provided more varied responses as to whether line manager approval was a pre-requisite for enrolling on programmes; suggesting that some practices are implemented locally rather than across

the institutions. None of the participants demonstrated awareness of what happens to the evaluation data once it had been collected but more than half of the participants indicated that they would like to know how the trainer or organisation had responded to their feedback.

The format, method and timing of collecting evaluation data

A common thread across participants was that customised-evaluation forms should be used and that this would encourage participants to spend more time on them, supporting the research of (Lingham, Richley and Rezania, 2006). One participant from Focus Group Three was particularly aggrieved at just having to 'ignore' questions that are irrelevant on the generic evaluation forms "If [the form was changed], the response rate would go up".

Whilst the phrase 'happy sheets' was not used by author A in the focus groups, the participants commonly aligned with the concept driven from Kirkpatrick's (1977) model of evaluation at level one, 'reaction'. For example referring to *'enjoying'* the session, *'liking'* the trainer, *'feeling anxious'* about working in mixed groups and finding the presentation slides *'easy to read'*. When prompted, some participants spoke of using extended questionnaires which required them to give examples of how the intervention could impact on their job role; thus providing evidence of evaluation at level two of Kirkpatrick's (1977) model.

Mixed attitudes were shown towards having to complete evaluation forms at the end of each training session, with participants stating that they needed the opportunity to reflect upon what they had learned. Non-academic participants were most likely to offer the view that the quality of the feedback was less considered when it had to be provided instantaneously:

Participant F: "You don't put the truth, just what people want to hear"

Participant G: "The tendency may be to be over-positive as they [the questionnaires] are immediate"

Regardless of whether or not an 'end of session evaluation' is carried out, all participants across both Universities raised the issue (without prompting) of a form of evaluation at a later date. In support of the issues identified by Chappell (2008), Participant C commented "Three months evaluation is a really good idea, it prompts you". Conversely, Participant D exclaimed *"I hate it, I have too much to do and can't be bothered"*. These two comments highlight the dichotomy of participant responses across focus groups when the issue of delayed evaluation were debated, interestingly there were no differences between the academic and non-academic groups. Collecting evaluation data after employees have returned to the workplace would enable organisations to undertake a higher level of evaluation; addressing some of the questions posed at level three of Kirkpatrick's (1977) model. However, Participant H (an academic) suggested: "those with strong views are most likely to return forms" which infers that the overall feedback may not be representative of the group of learners.

A range of suggestions for ways in which evaluation could be improved were offered by participants. Some were of a practical nature, such as the need to use plain English within the documentation (this was emphasised strongly by the non-academic participants) and better use of IT systems to provide feedback. There were diverse responses to the idea of incorporating 'testing' elements into the evaluation process, e.g., to evidence how behaviours/skills/knowledge have developed as an output of training interventions, the non-academic groups had the strongest feelings on this subject (Groups 1,2,3 and 4). Consensus was reached within all focus groups that they should only be used in limited circumstances, for example in IT programmes. This would appear to

be a positive result for both Universities, as the earlier work of Mann (1996) inferred that these techniques are not usually warmly received. Non-academics at University X also suggested that the phrase 'testing' should not be used at all, as it may raise barriers.

Multiple participants across the focus groups commented that evaluation feels like a *'rushed process'* and one mainly used to meet the needs of the HR Department. There were also suggestions from the academic groups that the trainers delivering the sessions did not appear to have *'bought into'* the importance of evaluation. Or if they had, this was not transmitted positively to the learners. Participant I extended this further by highlighting the need for space to reflect on the learning before completing an evaluation.

Whilst the work of James and Roffe (2000) suggests that it may be difficult to get learners to engage in providing qualitative feedback, the focus group participants indicated they could see the benefits of this (and would be happy to participate) in appropriate situations, such as at the conclusion of a multi-session intervention. Four non-academic participants from University Y shared a positive experience of an evaluation interview at the end of an NVQ programme, and their perception was that they provided more detailed feedback than would normally have been captured "It was good to talk about it and I could give her examples to answer her questions… I think I was on the phone for about half an hour".

Each focus group was asked how they felt that the evaluation process could be further improved as Broad (1997) emphasised the importance of involving learners in the design stages). Academic members from Group 7 felt that lessons could be learned from how students' training is evaluated and that this could be translated into a staff development context. Their academic peers in Group 8 highlighted the need to provide more detailed evaluation mechanisms at the end of extended projects/interventions. The non-academic participants in Groups 1 and 2 focused their recommendations around general communication issues, rather than concentrating specifically on the evaluation stage.

The role of line managers in the evaluation process

There is evidence from the focus groups of line managers going through the evaluation results with members of their team, thereby responding to one of the earlier criticisms of Kirkpatrick, (1998). This was most likely to happen where individual Departments had developed their own policy for carrying this out, rather than as a result of a University requirement. A number of the non-academic participants in University X described potential bottlenecks where they are accountable to both a supervisor and line manager, with discussions about development and evaluation falling between the two.

Without specifically being asked, non-academic participants raised ideas about how line managers could provide enhanced support e.g., provision of time for thinking and reflection and a more formal evaluation of training within annual performance reviews. Participants also suggested that if line managers received a copy of the evaluation forms, or engaged in a conversation with those who had attended, then they could make more informed decisions about future participation (which Griffin, 2011 identified as a key purpose of evaluating training). Non-academic staff participants in Groups 3 and 4 suggested evaluation could be a fixed item on monthly meeting agendas to share experiences and inform future plans.

Participants from Group Four suggested diverse practices between different line managers, with only some being proactive in arranging evaluation meetings with their team members. Those who

had attended were positive about its impact and felt it provided the chance to identify ways they could operationalise some of their new knowledge/skills/behaviours. This was not an unexpected response as the challenge of consistency between first line managers is an on-going HR issue (Author A and Author B, 2011). One focus group member put forward an interesting example "Some managers have a lot of evaluation data and don't know what to do with it. They have 'evidence' and would like to do something with it". Perhaps this also reflects the importance of engaging line managers at an earlier stage of the training cycle.

Discussion

The focus group participants provided reflections and suggestions consistent with 'best practice' outlined in the literature. For example using customised evaluation tools (Lingham *et al*., 2006) and leaving a gap between the date of intervention and collecting evaluation data (Chappell, 2008). One notable exception was the positive feelings towards providing feedback qualitatively which contradicts the work of James and Roffe (2000), who suggested engagement is usually difficult using these techniques.

Whilst Campbell's (1998) research clearly put responsibility for designing evaluation mechanisms to trainers, an overall observation from our findings is that the learners' were motivated to be involved in evaluation and identified useful views and ideas. An area of improved practice in university evaluation is to extend engagement of learners in evaluation, therefore implementing the recommendations from Broad (1997) to reflect the importance of the learners as key stakeholders and incorporate their insights and reflections.

The question of what happens to evaluation data was identified in the extant literature, with Pineda (2010:691 stating "the distribution and use of the results is what really gives meaning and value to the evaluation of training, enhancing the active involvement of all concerned" and this was reflected in our findings. The findings suggest evaluation does take place on a regular basis and there was some evidence of moving towards innovative practices and away from evaluation 'happy sheets'. It was apparent from the focus groups that employees undertaking training interventions in both institutions would benefit from having a greater understanding of evaluation and why it is being carried out. As a consequence of the findings we can confirm support of Anderson's (2009) assertion that a 'one size fits all' approach is unlikely to be appropriate.

Following analysis of the focus group data, it is apparent that the role of trainers in evaluation appears ever-expanding in terms of learners' expectations. The role of the line manager is important and we argue that this needs to go beyond the role of trainers. Pineda (2010) confirms that line managers are key stakeholders who can make a difference and this is consistent with the wider line manager literature, for example the recent work of (Author A and Author B, 2011). Findings from the focus group clearly illustrated pockets of good practice in the role played by line managers in supporting (or not) their team members. The mixed experiences reported suggest that *consistency* remains an organisational issue.

Informing the development of the training evaluation toolkit

Significantly, the involvement of the two University HR Departments and staff participants in the focus groups in the research raised consciousness to the importance of evaluation of training and development and prompted reflections on current practice and the potential for improvements in practice. As part of a funded project it was critical that the findings of the research informed the design of the evaluation tools within the practical evaluation toolkit and this can be seen within the

(free) project website available at:
http://www.northumbria.ac.uk/sd/central/hr/peopledevelopment/evaluation/hefce/.

Due to the high level of organisational expenditure dedicated to training and development interventions (UKCES, 2012) there is a need for a more holistic approach to the way in which evaluation frameworks are planned, implemented and followed up. Benefits of updating evaluation processes include; more valuable data which can be used for decision making, increased accountability for trainers, improved alignment with organisational goals and objectives and greater opportunities to meet the specific needs of the learners and support them in transferring their learning back into the workplace. The training evaluation toolkit we developed for UK universities is offered as a means by which HEIs can engage in developing a more holistic approach to evaluation.

Conclusions, limitations and implications for further research

 In exploring our guiding research question, what are staff perceptions of how their training and development is evaluated in two UK HEIs? We contend that it is important to consider the views of an important stakeholder group; the learners. We achieved this by using the medium of focus groups.

In summary, this study offers new knowledge to the HRD field by presenting research into training evaluation utilising qualitative methods in the Higher Education context. In doing so we have addressed a need for HE evaluation research identified by Burgoyne *et al*. (2009); the need for qualitative evaluation research identified by Diamantidis and Chatzoglou (2012) and have valued the views of an important stakeholder group; the learners, via their involvement in focus groups. In engaging learners in the research we make a key contribution to the HRD research and practice. The alignment between the core issues of evaluation identified by the interviewees in HEIs and the more general extant evaluation literature suggests that the findings are applicable beyond the HE sector.

Limitations

The focus on two UK University case studies and engagement of 47 learners are the most obvious limitations to this study. Future research could replicate the study in other universities and geographical regions. As outlined in an earlier study by Butler (2010) a further potential limitation of using focus groups is that individual voices can get lost within the group. In addition, participants rely on their memories of past evaluation experiences, however this is a key characteristic of qualitative research. Also our emphasis was formal training interventions which may only be one element of participants' development journeys.

Implications for further research

This paper outlines an exploratory study and focuses on the experiences of a small sample of learners, replicating this approach on a larger scale (and perhaps with the use of some quantitative tools) would progress the field. A recurring theme in this study is the role of trainers and how their role appears to have increased in evaluation. Therefore the knowledge base could be extended through future research in this area; undertaking primary research with trainers would address the current gap. A further contribution could be achieved through the use of a longitudinal research design which tracks learners throughout a learning journey and captures their experience of the different evaluation mechanisms.

Recommendations

A range of recommendations stem from the findings of this study. Based on the primary and secondary research it is apparent that Trainers should ensure their learners understand reasons for evaluation (Campbell, 1998) and establish what level of feedback is needed for different interventions, our participants were clear that a one size fits all approach is not appropriate. Furthermore, in line with the earlier work of Diamantidis and Chatzoglou (2012) we recommend that evaluation methods should be considered at the design stage and incorporate more diverse data collection methods. This was reinforced in our focus groups where the participants recognised the value of different evaluation methods including those of a qualitative nature.

It is clear from both the primary and secondary data that questionnaires continue to be the most frequently used tool, however we recommend that differentiated evaluation questionnaires should be used, reflecting the different types of interventions. Whilst open-ended questions are useful; the potential negative impact on response rates should be considered. In line with the work of Pineda (2010) we also subscribe to the view that organisations' need to undertake a critical and transparent review of the evaluation data should take place by someone suitably qualified.

Finally, our focus group findings reinforced the extant literature view of the crucial role of first line managers in the evaluation process and participants gave several examples of inconsistencies in how different line managers provides support. We therefore recommend that first-line managers should receive on-going support to understand their role in supporting their employees' training and development through contributing to the evaluation process.

Implications for practice

Based on the primary research, a key implication for practice is the need to consider evaluation as soon as new learning programmes are being designed or commissioned (ideally with input from the prospective learners, Broad, 1997) and the importance of sharing anticipated learning objectives with the learners. Our participants recognised the usefulness of having personal learning objectives to measure their progress against. This was also reinforced in the learners' assertion that those attending the events should take some personal responsibility. On a very practical level our participants emphasised the importance of using plain English within evaluation mechanisms.

Through the primary and secondary data it is clear that organisations are best served through the use of a menu of evaluation tools so that the most appropriate can be selected in terms of the level and amount of information required. As questionnaires continue to be the most frequently used mechanism, organisations need to be able to customise them to meet the specific needs of events, trainers and learners. The significance of evaluation timing was also illuminated by our research with participants recognising a need for time for reflection and a follow up evaluation stage after they have had the opportunity to put their learning into practice.

Finally, a clear implication for practice that was emphasised throughout our focus groups was the vital role of Trainers in the way that evaluation is presented and explained to the learners. Our participants believe that Trainers need to be seen to buy into the process and avoid it being a rushed afterthought at the end of an event. Furthermore, our learners were keen for the feedback loop to be closed by the sharing of outcomes that occur as a consequence of their feedback.

References

Alliger, G.M. and Janak, E.A. (1989), "Kirkpatrick's levels of training criteria: Thirty years later", *Personnel Psychology*, Vol. 42 No. 2, pp. 331-342.

Alvarez, K., Salas, E. and Garofano, C.M. (2004), "An integrated model of training evaluation and effectiveness", *Human Resource Development Review,* Vol. 3 No.4. pp. 385-416.

Anderson, V. (2009), *The value of learning.* CIPD, London.

APSC (2005), *Evaluating learning and development: A framework for judging success.* Australian Public Service Commission, Canberra.

Barnett, R. (2003), *Beyond reason: Living with Ideology in University.* Open University Press, Buckingham.

Bimpitsos, C. and Petridou, E. (2012), "A transdisciplinary approach to training: Preliminary research findings based on a case analysis", *European Journal of Training and Development,* Vol. 36 No. 9, pp. 911-929.

Blackmore, P. and Blackwell, R. (2004), *How can staff development be more strategic?* Leadership Foundation for Higher Education, London.

Bramley, P. and Kitson, B. (1994), "Evaluating Training against Business Criteria", *Journal of European Industrial Training,* Vol. 18 No. 1, pp. 10-14.

Broad, M.L. (1997), "Overview of transfer of training: From learning to performance", *Performance Improvement Quarterly,* Vol. 10 No. 2, pp. 7-21.

Burgoyne, J., Mackness, J. and Williams, S. (2009), *Baseline study of Leadership Development in Higher Education.* Leadership Foundation for Higher Education, London.

Butler, M.J.R. and Reddy, P. (2010), "Developing critical understanding in HRM students: Using innovative teaching methods to encourage deep approaches to study", *Journal of European Industrial Training,* Vol. 34 No. 8/9, pp. 772 – 789.

Campbell, C.P. (1998), "Training course/program evaluation: Principles and practice", *Journal of European Industrial Training,* Vol. 22 No. 8, pp. 323-244.

Cassell, C. (2012),'Exploring dignity at work using three types of qualitative data analysis: themes, metaphors and stories', paper presented to Research Methodology track, *The British Academy of Management Annual Conference, 11-13 September*, University of Cardiff.

Cefai, J. (2009), "Evaluation of staff learning and development at Anglia Ruskin University Library", *SCONUL*, Vol. 47, pp. 42-44.

Chiaburu, D.S. and Lindsay, D.R. (2008), "Can do or will do? The importance of self-efficacy and instrumentality for training transfer", *Human Resource Development International,* Vol. 11 No. 2 pp. 199-206.

CIPD (2012), *Learning and talent development survey.* CIPD, London.

Darby, J.A. (2006), "Evaluating training courses: an exercise in social desirability?" *Journal of European Industrial Training,* Vol. 30 No. 3, pp. 227-239.

Darby, J.A. (2007), "Open-ended course evaluations: A response rate problem?" *Journal of European Industrial Training,* Vol. 31 No. 5, pp. 402-412.

Dawson, R. (1995), "Fill this in before you go: under-utilised evaluation sheets", *Journal of European Industrial Training,* Vol. 19 No. 2, pp. 3-7.

Diamantidis, A.D. and Chatzoglou, P.D. (2012), "Evaluation of formal training programmes in Greek Organisations", *European Journal of Training and Development,* Vol. 36 No. 9, pp. 888-910.

Dyer, S. (1994), "Kirkpatrick's Mirror", *Journal of European and Industrial Training,* Vol. 18 No. 5, pp. 31-32.

Gordon, G. and Whitchurch, C. (2007), "Managing Human Resources in Higher Education: The implications of a diversifying workforce", *Higher Education Management and Policy,* Vol. 19 No. 2, pp. 135-155.

Griffin, R. (2011), "Seeing the word for the trees: Workplace learning evaluation", *Journal of European Industrial Training,* Vol. 35 No. 8, pp. 841-850.

Gubbins, C., Corrigan, S., Garavan, T.N., O'Connor, C., Leahy, D., Long, D. and Murphy. E. (2012), "Evaluating a tacit knowledge sharing initiatives: A case study", *European Journal of Training and Development,* Vol. 38 No. 8, pp. 827-847.

Hamblin, A.C. (1974), *Evaluation and Control of Training.* McGraw-Hill, New York

HESA (2012),"Statistics: Staff employed at UK HE Institutions", available at: http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=1898&Itemid=239 (accessed 1 March 2013).

James, C. and Roffe, I. (2000), "The evaluation of goal and goal-free training innovation", *Journal of European Industrial Training,* Vol. 24 No. 1, pp. 12-20.

King, N. (2004), *Using templates in the thematic analysis of texts'.* In Cassell, C.M. and Symon, G. (eds), *Essential Guide to Qualitative Methods in Organizational Research,* pp. 256-270. Sage, London.

King, N. (2012), *Template analysis'.* In Symon, G. and Cassell, C.M. (eds), *Qualitative Organizational Research: Key Challenges and Core Methods*, pp. 426-450. Sage, London.

Kirkpatrick, D.L. (1977), "Evaluating training programs: Evidence vs proof", *Training and Development Journal,* Vol. 31 No. 11, pp. 9-12.

Kirkpatrick, D.L. (1998), *Evaluating training programs: The four levels.* 2nd edition. Berrett-Koehler Publishers Inc, San Francisco.

Lewis, P. and Thornhill, A. (1994), "The evaluation of training: An Organizational Approach"*, Journal of European Industrial Training,* Vol. 18 No. 8, pp. 25-33.

Lingham, T., Richley, B. and Rezania, D. (2006), "An evaluation system for training programs: A case study using a four phase approach", *Career Development International,* Vol. 11 No. 4, pp. 334-351.

Mann, S. (1996), "What should training evaluations evaluate?" *Journal of European Industrial Training,* Vol. 20 No. 9, pp. 14-20.

McGuire, D., Garavan, T.N., O'Donnell, D., Saha, S.K. and Cseh, M. (2008), "Managers' personal values as predictors of importance attached to training and development: a cross country exploratory study", *Human Resource Development International,* Vol. 11 No. 4 pp. 335-350.

McGivern, Y. (2009), *The Practice of Market Research,* 3rd edition. Pearson Education, Essex.

McLelland, S.B. (1994), "Training needs assessment data-gathering methods: Part 3, Focus Groups", *Journal of European and Industrial Training,* Vol. 18 No. 3 pp. 29-32.

NTSF (2000), *Third Report of the National Skills Taskforce: Tackling the Adult Skills Gap: Upskilling Adults and the Role of Workplace Learning.* DfEE, London.

Phillips, J.J. (2002), *Return on Investment in Training and Performance Improvement Programs,* 2nd edition. Butterworth-Heinemann, Woburn.

Pineda, P. (2010), "Evaluation of training in organisations: A proposal for an integrated model", *Journal of European and Industrial Training,* Vol. 34 No. 7, pp. 673-693.

Russ-Eft, D. and Preskill, H. (2008), "Improving the quality of evaluation participation: A meta-evaluation", *Human Resource Development International,* Vol. 11 No. 1, pp. 35-50.

Sambrook, S. (2001), "Factors influencing learners' perceptions of the quality of computer based learning materials", *Journal of European Industrial Training,* Vol. 25 No. 2/3/4, pp. 157 – 167.

Short, T. (2009), "Exploring the vacuum in training evaluation: Is this a case of mission impossible?", *Development and Learning in Organizations,* Vol. 23 No. 5, pp. 15-18.

Tennant,C.; Boonkrong, M. and Roberts, P. (2002), "The design of a Training Programme Measurement Model", *Journal of European Industrial Training*, Vol. 26 No. 5, pp. 230-240.

Towler, A.J. and Dipboye, R.L. (2009), "Effects of Trainer Expressiveness, Organization and Training Orientation on Training Outcomes", *Journal of Applied Psychology,* Vol. 86 No. 4, pp. 664-673.

UKCES (2012), *UK Employer skills survey 2011: First findings.* UK Commission for Employment and Skills, London.

Warr, P., Bird, M. and Rackham, N. (1970), *Evaluation of management training: A practical framework, with cases, for evaluating training needs and results.* Gower, London

Weissner, C.A., Hatcher, T., Chapman, D. and Storberg-Walker, J. (2008), "Creating new learning at professional conferences: an innovative approach to conference learning, knowledge construction and evaluation", *Human Resource Development International,* Vol. 11 No. 4, pp. 367-383.

Whitchurch, C. (2008), *Professional Managers in UK Higher Education: Preparing for complex futures.* Leadership Foundation for Higher Education, London.

Table 1. Overview of Kirkpatrick's model of learning evaluation

Level One	Reaction	How do the learners feel about the programme? To what extent are they 'satisfied customers'?
Level Two	Learning	To what extent have the learners learned the information and skills? To what extent have their attitudes changed?
Level Three	Behaviour	To what extent has their job behaviour changed?
Level Four	Results	To what extent have [the learners] results been affected?

Source: Kirkpatrick, D.L. (1977) 'Evaluating training programs: Evidence vs proof',

Table 2. Characteristics of focus group participants

Focus group	Participants	Job role characteristics	Training intervention attended 12 months prior to the focus group
Focus Group 1 University X	3 female 4 male	Non-academic staff Manual workers	Mandatory event e.g. University health and safety
Focus Group 2 University Y	5 female 3 male		
Focus Group 3 University X	4 female 1 male	Non-academic staff Administrative workers	Supervisory skills
Focus Group 4 University Y	2 female 2 male		
Focus Group 5 University X	4 female 1 male	Academic staff	Leadership
Focus Group 6 University Y	3 female 3 male		
Focus Group 7 University X	4 female 2 male	Academic staff	Learning & teaching OR research
Focus Group 8 University Y	4 female 2 male		

Table 3. Key focus group questions for evaluation of training interventions

- How important do you think evaluation is? Why?

- Why do you think you are asked to complete an evaluation form?

- How do you feel about the current evaluation forms? Why?

- What usually happens when you return to the workplace?

- What are the opportunities to transfer learning into the workplace? What are the key issues?

- How often are you asked to undertake a second evaluation?

- How comfortable would you feel with some of the alternative ways that evaluation can be carried out?

- How would you feel about having to do 'before and after' tests before you could attend a training event?

- What do you think are the barriers to effective training (and evaluation of training) are?

Table 4. Examples of data analysis

Data Example	First Order Theme	Second Order Theme
I think the trainer should know how they are performing. I know there are different levels so they (trainers) need to tailor to the audience and it needs to be passed back.	Reason for evaluating training	The role of the trainer/facilitator. How evaluation data is used.
Three months evaluation is a really good idea, it prompts you.	Format of collecting data	Having the opportunity to use the learning i.e. skills, behaviours, knowledge
I personally don't pass to the line manager, I feel my training is my responsibility.	Role of the line manager	Roles of different stakeholders in the training and evaluation process
Yes, if changed [the evaluation form] the response rate would go up.	Format of collecting data	Customising evaluation tools so that they are appropriate for different training interventions.