The Role of Evaluation in Planning and Learning from Short Training Programs

Pamela St. Leger
Visiting Professor Department of Learning Science
Graduate School of Education Hiroshima University
and
Centre for Program Evaluation Melbourne
Graduate School of Education The University of Melbourne Australia
pksl@unimelb.edu.au

Abstract

This paper reviews the role of short course training programs within broader organisational learning frameworks and the part that evaluation plays in informing decision making about future development.

Four dimensions of program planning and evaluation are discussed. These are: (a) the kinds of learning that can reasonably be expected from short course training programs; (b) approaches and methods of evaluation that may be used to gain information about the value of short course training interventions; (c) other aspects of training and workplace learning that might be added to short course interventions to increase participants' use of knowledge and skills gained; and given this broader concept of workplace learning, (d) evaluation approaches that might be used to provide useful information to program managers to aid decision making about future training needs and program provision.

While short training courses may be easier to organize than comprehensive and systematic training and development programs, they may have little effect on changing workplace practices. In an increasing organisational climate of accountability for use of resources for which there are competing demands, coupled with organisational goals to achieve strategic advantage over competitors, short courses need to be located within broader strategic plans to develop organisational learning capacity.

Keywords

Training and Development, Training Program Design, Training Evaluation,
Transfer of Training, Capacity Building

1. Introduction

Short seminars and courses have long been a prominent feature of training programs in a wide range of sectors that includes business, education, health and community development. Usually, program managers attach an evaluation activity of some kind to "get feedback" about the value of the session or course to the participants. This information may then be used for a range of purposes: for example, to justify spending and allocation of other resources and/or the continuation of the course, or to inform the design of subsequent programs. However, stand alone short course programs and limiting evaluation to participant feedback present a narrow view of training, learning and development.

This paper discusses: (a) the kinds of learning that can reasonably be expected from short course training programs; (b) approaches and methods of evaluation that may be used to gain information about the value of short course training interventions; (c) other aspects of training and workplace learning that might be added to short course interventions to increase participants' use of knowledge and skills gained; and given this broader concept of workplace learning, (d) the evaluation approaches that might be used to provide useful information to program managers to aid decision making about future training needs and program provision.

2. What Kinds of Learning can be Reasonably Expected from Short Course Training Programs?

Before embarking on a discussion of short course training programs, I want to make clear that within the context of current thinking and trends in training and development, such a narrow focus would be regarded as somewhat outdated. This is because knowledge of what does and what does not constitute effective training has shifted in the last decade or so from a focus on training individuals to a focus on developing organisational learning through systemic planning to achieve business and organisational goals (Rohmetra and Easter-by Smith 2004). Thus, short courses are now regarded as one intervention in the training and development process that complement, or are complemented by, other interventions such as ongoing mentoring and coaching, and management practices that promote and support organisational learning.

Therefore, due consideration needs to be given to the kinds of learning that result from various training activities, how and when they are scheduled, and the conditions that are needed to support learning. For example, Tannenbaum and Yukl's (1992) Analysis of studies of effectiveness of simulation games used in management education concluded that mechanisms such as debriefing should be built into simulation training activities. They also concluded that simulations alone are not likely to change management behaviour. Additional complementary activities such as "opportunities for coaching, feedback, reinforcement and practice of skills on the job" (p. 408), and a clearly articulated model of effective and appropriate management were necessary to produce positive training results. Organisational approaches to training and building organisational capacity are discussed in Sections 4 and 5 of this paper.

With respect to selecting training methods that are likely to produce different kinds of learning, Rothwell & Sredl (1992) offer a useful planning framework. They suggest the first step is planning an "organisational curriculum". This is "a long-term, strategic instructional plan for all formal learning events" (p.341) which provides "long-term direction" for training to assist an organisation to:
  * adapt to change; and
  * maintain effective and efficient work practices.
Thus the "curriculum" is both a "policy" and a "tool" to guide learning. Within this framework, Rothwell and Sredl (1992) make explicit the links between particular theories of instruction and different types of organisational goals and purposes. These links are outlined below:

Table 1 Theories of Instruction matched to Organisational Goals and Purposes

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Theory of Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>To disseminate new information</td>
<td>Subject centred curriculum</td>
</tr>
<tr>
<td>To develop technical skills</td>
<td>Objectives centred curriculum</td>
</tr>
<tr>
<td>To generate new ideas</td>
<td>Experience centred curriculum</td>
</tr>
<tr>
<td>To match learners' needs to relevant learning experiences?</td>
<td>Opportunity centred curriculum</td>
</tr>
</tbody>
</table>

Source: Rothwell and Sredl (1992)

Kay (1995) uses similar logic to identify particular training methods in relation to four categories of purpose. These are outlined in Figure 2 below:

Figure 2 Which Training Methods? For What Purposes?

Reflective Nature of participation and involvement of students Active
These kinds of training logic frameworks help to make decision making about what, why and how to evaluate training activities easier for both program managers and evaluators because the clarity of purpose and plausibility of training program designs that are created from them delivering desired outcomes are likely to be clearer. Many evaluators now regard clarification of the links between purpose, design and intended outcomes as the first step in any evaluation (Owen, 2006). Without such coherence, it is difficult to attribute program effectiveness (or ineffectiveness) with confidence. Program clarification evaluation (program theory building) is discussed later in this paper.

3. Approaches and Methods of Evaluation that may be used to Gain Information about the Value of Short Course Training Interventions

Donald Kirkpatrick's (1998) Model of Evaluating Training is well known all over the world. It is a useful framework to apply to the logic of short course training activities because it helps program managers to recognise which dimensions of training outcomes can be realistically evaluated given the nature of the training activities and what it is realistic for participants to learn within those activities. Thus, they can direct scarce evaluation resources (ie. money and personnel time) to gathering useful data to inform decision making about ongoing program development rather than simply engaging in symbolic evaluation to demonstrate compliance.

Kirkpatrick (1998) identifies four levels of evaluation to measure training effectiveness. These are:
- Level 1 Reaction: What participants liked and/or disliked about the training;
- Level 2 Learning: Knowledge, skills and shifts in attitudes participants gained from the training;
- Level 3 Behaviour: Knowledge, skills and attitudes participants apply routinely in their work; and
- Level 4 Results: Impacts on the organisation, for example, reduced absenteeism, increased productivity.

The first two levels of Kirkpatrick's model have been widely used evaluate training effectiveness. The use of simple self-reporting instruments such as reaction surveys is common. These range in format from simple check box and Likert scale responses to item statements to questions that require participants to write more descriptive responses. The extent to which they are of value depends on the skill and experience of the person who designed the survey items and his or her skills in analysis and reporting findings, and the relevance of the information produced to the ultimate user/s. Kirkpatrick points out that the kind of information that they generate is of limited use in making decisions about future program provision. However, he does claim that a measure of participants’ reaction to a training experience may be important to gauge participant levels of motivation to undertake a particular form of training. This may be an important consideration for e-learning training programs where participants' success depends so much on their comfort with the medium of delivery. For example, Strother (2002) found that measures of reaction provide useful evidence about levels of motivation and satisfaction with the learning format in e-learning training programs.

However, the extent to which participants liked the experience they had in a training course may imply that it was useful to them, but this level of evaluation but it does not claim to produce evidence that participants are likely to be able to apply what they learned in their work. Tannenbaum and Yukl (1992) argue that a better predictor of the worth of a training intervention may be to ask participants to focus their reactions on how useful and how applicable the training was to their work situations rather than whether or not they liked the training.

Kirkpatrick advocates the use of pre and post tests or control groups to evaluate Level 2: Learning. These kinds of measures, he claims, provide objective information as “proof” of increased learning. However, Kraiger,
McLinden and Casper (2004) caution that due consideration should be given to the extent to which training can be isolated from other factors that may affect work performance. For example, increased sales may be due in part to a sales training intervention, but they may also be due to parallel marketing activities. To counteract any devaluing of training contributions to business results, Kraiger et al. (2004) propose comparing sales results between a group which received training and a control group which did not. Evaluation designs of this kind, they argue, “can greatly enhance the evidence that changes in criteria are due to training and not alternative factors” (p. 345). Nevertheless, in employing such designs, evaluators also need to consider any potential ethical issues that might arise from withholding a training intervention to individuals who make up a control group.

However, Brinkerhoff (2003) argues that it is important to identify different contextual factors that may influence participants’ take up of training. In order to do this, he advocates use of his Success Case Method which involves selecting most successful and least successful respondents from survey data, and interviewing them to develop best and worst case scenarios or stories of what worked, or did not work, and why. This evaluation approach relies on “evidence” of participants' experiences to seek understanding of a training initiative rather than “proof” that it did, or did not work.

Such an approach may also capture antecedents such as participants’ existing expertise and intrinsic motivation to undertake training. For example, Green and Skinner (2006) found that their evaluation of a time management short course revealed that, in one instance, a group of senior management consultants showed “modest improvement” in comparison to a “younger, less experienced group”, the difference being that the former group were already “highly skilled in most aspects of management” (p. 135).

4. Other Dimensions of Training and Workplace Learning

Kirkpatrick (1998) maintains that, for participants’ learning from short course activities to be transferred into change in their routine work practice, Level 3: Behaviour, four conditions are necessary. The first two relate to motivation and knowledge and skill acquisition, that is, (1) participants must want to change their work practices and the “dose” (to use a health metaphor) and quality of the training intervention needs to be sufficient for (2) participants to know what do and how to do it. Green and Skinner (2006) have identified another layer of complexity when "behaviour" is applied to organisational or personal management concepts such as time management. Citing Atkinson (1999), they point out that “soft skills” are based on:

\[\text{cognitive skills; thought processes which allow the individual to read situations and which can be used to understand and address issues or problems, the inherent assumption being that individual effectiveness is inextricably linked to organisational effectiveness} \] (p. 135).

Thus, Kirkpatrick’s remaining conditions for changing practice, that is, (3) a supportive work climate and (4) incentive for changing work practice need to be taken into consideration. Regardless of how good a short course training program is, or how much participants’ liked it, decisions that affect the degree to which participants may change work practices may be outside the control of training personnel (Tovey and Lawlor 2004) or the participants. For example, managers need to be willing to provide opportunities for participants to put new knowledge and skills into practice immediately on their return, and possibly during their training program, along with coaching and mentoring support. Also, in some cases, incentives for undertaking training may have more to do with increases in salary rather than intrinsic motivational reasons. At the other end of the training process, lack of transfer may be
affected by other organisational factors such as increased workloads due to reduced staffing (Green and Skinner 2006).

Much of the recent literature on training and development, and evaluation of it, focuses on the transfer issue and how training policies, interventions and necessary supportive conditions and infrastructure can be built into organisational strategic planning (Caffarella 2002; Foley 2001; Haskell 1998; Kraiger et al. 2004; Salas and Cannon-Bowers 2001; Tennant 1999; Tovey and Lawlor 2004).

5. Evaluation for Decision Making about Future Training Needs and Program Provision

The shift in focus in the training and development literature from individual to organisational learning has also turned attention to human resource capacity building, accountability for time and money allocated to training and development activities, and results (Caffarella 2002; Foley 2001; Inman, Blumenfeld, and Ko 2005; Kraiger et al. 2004; Rohmetra and Easterby-Smith 2004; Salas and Cannon-Bowers 2001; Tovey and Lawlor 2004). Rohmetra and Easterby-Smith (2004), citing the banking industry, argue that, in addition to business strategies such as “repositioning” institutions through acquiring other companies, and developing more distinctive products to be more competitive, organisations also recognize the need to review operational functions such as training and development. Thus this function is now viewed through a more systematic lens.

As such, a more systematic approach to evaluation is required. Over a decade ago, Tannenbaum and Yukl (1992) concluded that the focus of research on training effectiveness required a shift from training methods to articulating the nature and purpose of training interventions, and what works in particular situations and contexts, why and for whom. Making explicit, firstly, the purpose of training and, secondly, the causal links between planned activities of the intervention and desired outcomes to determine its plausibility. This kind of evaluation approach (known as program theory building or program logic) is gaining increasing acceptance in sectors such as international development, health and community development as they strive to better understand how interventions designed to build capacity work in particular settings.

However, the literature indicates that the training sector has not adopted this kind of thinking to such an extent, for example, Strother (2002) cites a American Society of Training and Development (ASTD) study (2000) that “found that while 95 percent of surveyed organisations gauged trainees’ reactions to courses (e.g. how well they liked the courses), ... only three percent of respondents [in that study] make a real effort to measure the business results of [their net-based] training programs” (p. 3). Strother also noted that reviews of previous research (Perraton 2000; Saba 2000) indicated that few research and evaluation studies of distance education programs were based on sound methodological frameworks. They reported that “the researchers simply carried out their experiments in which they compared distance learning with traditional classroom content delivery and reported statistical results” (p. 8).

Such simple measures fail to take into account the teaching requirements of different learning formats such as e-learning, mentoring and coaching, and other personal and external variables that might influence participants’ outcomes. Thus, there appear to be large leaps of faith in the internal coherence between the purposes, content, activities and outcomes of these kinds of programs, let alone the effects of local factors that might contribute to different levels of success (or failure) at particular sites.

In order to take these complexities into account, many evaluation theorists (Davidson 2005; Funnell 2000; Owen 2006; Patton 1996; Pawson and Tilley 1997; Rogers, Petrosino, Huebner, and Hacsi 2000) advocate utilization-focused approaches such as needs assessment (organisational and participant), program theory development (to identify the critical components of training programs), and planning and implementing evaluations...
in close consultation with program managers and other key users of evaluation findings.

It is only when this critical clarificative evaluation work is done that it might be realistic to undertake an impact study of a training program or what Kirkpatrick refers to as Level 4: Results. However, it is difficult to prove that a training intervention has contributed to organisational effectiveness. What participants choose or choose not to embrace from a training experience is entangled with personal antecedents and the organisational climate that surrounds them. Also, from a practical perspective, Green and Skinner (2006) argue that:

For the average practitioner the identification of direct causal links in terms of impact on the organisation's 'bottom line' is often perceived as problematic, complex and daunting, particularly in relation to the development of soft skills and can act as a significant barrier to evaluation beyond the reaction level (Skinner 2004).

Impact studies are based on the assumptions that there is good documentation that clearly articulates the program's objectives (not common according to Owen, 2006) and that training programs have been running long enough to show effects over time or that participants and their organizations are followed up some months later to investigate effects. The scale and therefore the expense to undertake such extensive evaluation work may not be warranted (Phillips and Phillips 2001).

6. A Framework for Evaluating Training Programs

One way to determine the kind of evaluation that may be best suited to a particular set of circumstances is from a program development viewpoint, that is, at what stage of maturity is the training program, who is likely to use the evaluation findings, and for what purpose? These fundamental questions underpin Owen’s (2006) conceptualization of five forms of evaluation (proactive, clarificative, interactive, monitoring and impact). However, though working with education and training practitioners who are new to evaluation over many years, I have found a simplified adaptation of the model (see Figure 3) useful as a starting point in helping them to think evaluatively about their training programs (short course, mentoring or other learning and development activities) from the very beginning and within the context of broader organisational learning goals. Thus the evaluation framework can be used as a tool for informing decisions about both program design and tracking program effects at various stages of a program’s development.

Focusing evaluation on the proactive and responsive phases of training program development may also help to identify the supportive conditions and infrastructure that are critical to achieving the desired training outcomes. Once these elements are identified, Owen (2006) argues that there is a sound conceptual basis for then considering how other forms of evaluation such as monitoring and impact evaluation might be used to measure program effectiveness. The framework therefore presents a conceptual and practical view of evaluation work (that is, the gathering and analysis of evidence) as an integral part of policy and program development and decision making, and suggests the kinds of data collection techniques that could be used to provide useful evidence (Owen, 2006).

7. Conclusion

While short training courses may be easier to organize than comprehensive and systematic training and development
To counteract potential negative arguments about the relative value of training evaluation such as the cost, time required, and difficulty of attributing training initiatives to achieving business goals, training managers need to develop persuasive arguments to integrate evaluation into a framework that encapsulates all stages of program development. Used in this strategic way, evaluation resources may be used to greater effect to: review effective practice elsewhere; assess organisational and individual needs to develop training program theory; and refine program theory through action research projects generated by program staff. Once a clearly articulated and plausible program theory is developed, it should then be possible to design a monitoring system of evaluation that enables useful data to be collected to inform ongoing decisions about future training initiatives.

**References**


(Accepted 31st, 2008)