Springer
the language of science

Dear Author/Editor,

Here are the proofs of your chapter as well as the metadata sheets.

Metadata

- Please carefully proof read the metadata, above all the names and address.
- In case there were no abstracts for this book submitted with the manuscript, the first 10-15 lines of the first paragraph were taken. In case you want to replace these default abstracts, please submit new abstracts with your proof corrections.

Page proofs

- Please check the proofs and mark your corrections either by
  - entering your corrections online
    or
  - opening the PDF file in Adobe Acrobat and inserting your corrections using the tool "Comment and Markup"
    or
  - printing the file and marking corrections on hardcopy. Please mark all corrections in dark pen in the text and in the margin at least ¼" (6 mm) from the edge.
- You can upload your annotated PDF file or your corrected printout on our Proofing Website. In case you are not able to scan the printout , send us the corrected pages via fax.
- Please note that any changes at this stage are limited to typographical errors and serious errors of fact.
- If the figures were converted to black and white, please check that the quality of such figures is sufficient and that all references to color in any text discussing the figures is changed accordingly. If the quality of some figures is judged to be insufficient, please send an improved grayscale figure.

Metadata of the chapter that will be visualized online

Book Title		Transfer of Learning in Organizations
Chapter Title		The Measurement of Transfer Using Return on Investment
Copyright		Springer International Publishing Switzerland 2014
Corresponding Author	Family name	**Donovan**
	Particle	
	Given name	**Paul**
	Suffix	
	Division	
	Organization	National University of Ireland
	Address	Maynooth, Ireland
	email	Paul.Donovan@nuim.ie
Abstract		The transfer of training was, as a research area, originally focused on outcomes' evaluation in terms of reaction, learning, behavior, and results. Outcomes' evaluation, widely accepted by practitioners, is criticized by researchers seeking a more systemic approach for assessing the effectiveness of training interventions. As a result, the field of transfer research developed approaches more cognizant of context with a muted emphasis on outcomes. In turn, these approaches were criticized for their lack of tangible evidence of transfer of training. This chapter describes the development of evaluation from its early days up until its current evolution. The original outcomes' model, the Four-Level Model by Donald Kirkpatrick, is described and its contributions and criticisms are discussed. Phillips' return on investment approach (ROI) is also described. An account of how to measure the transfer of training using ROI is discussed. Finally, the implications of using ROI as a measure of transfer are considered.

Chapter 9
The Measurement of Transfer Using Return on Investment

Paul Donovan

9.1 Introduction

AQ1

The transfer of training, as an area of research has been developed out of the literature on evaluation of training. Originally, evaluation research was principally focused on outcomes from the learning process in terms of reaction, learning, behavior, and results and this is described in terms of content and process (Kirkpatrick 1959a). This approach, and the work of its adherent group, was roundly criticized by academics who sought a more holistic approach for the effectiveness of training interventions. Subsequently, some academics began to seek approaches for effectiveness which were more cognizant of context and process, and with a lesser focus on tangible outcomes (Baldwin and Ford 1988; Broad and Newstrom 1992; Holton 1996). These approaches concentrated more on creating measures of transfer that would indicate the effectiveness of training.

This chapter charts the development of evaluation from its early days up until its current evolution, as measurement of the transfer of training. It describes the early years and stages of development of transfer including its early atheoretical phase. Evaluation philosophy is discussed and note is made of the tendency toward objectivism and positivism in the approaches to evaluation and transfer. The original outcomes' model, the four-level model by Donald Kirkpatrick, is described and its derivatives, contributions, and criticisms are discussed. Phillips' ROI, also known as level 5 is also covered. A description of how to measure the transfer of training is discussed. Finally, the implications of using ROI as a measure of transfer are considered.

P. Donovan (✉)
National University of Ireland, Maynooth, Ireland
e-mail: Paul.Donovan@nuim.ie

K. Schneider (ed.), *Transfer of Learning in Organizations,*
DOI 10.1007/978-3-319-02093-8_9,

Chapter ID: 313740_1_En Chapter No: 9 Dispatch Date: 18-10-2013 Proof No: 1

9.2 Evaluation in Times Past

The evaluation of training and development interventions today is a development of early attempts to improve the process of education, particularly in the United States. In the early decades of the twentieth century, the popularity of the discipline of scientific management encouraged the measurement and assessment of people; American educators began to see the possibility of adopting these methods and applying them to educational improvement. By the 1920s a great deal of experimentation was taking place in educational establishments. It was decided that the greater availability of education to the masses and the greater range of abilities among pupils might require different approaches. In the United States, an evaluation program was set up to compare the traditional curricula with the more novel approaches. In order to make these comparisons it was decided to use the objectives of the educational approaches themselves as a means of evaluating those same approaches.

> The process of evaluation is essentially the process of determining to what extent the educational objectives are actually being realized …. however since educational objectives are essentially changes in human beings, that is, the objectives aimed at are to produce certain desirable changes in the behaviour patterns of the students, then evaluation is the process for determining the degree to which these changes in behaviour are actually taking place. (Tyler 1949, p. 105)

This approach was an advance over previous methods that focused on examination results and teacher's impressions of classroom work. Educational establishments understood and accepted the work of Tyler, especially the way it made explicit what they were trying to achieve.

9.2.1 Stages in Evaluation

An understanding of evaluation of training can be gained by tracing its development over the last half century. Wang and Spitzer (2005) suggest that the evolution of evaluation in human resource development (HRD) comprises three distinct stages: (a) practice-oriented atheoretical stage, (b) process-driven operational stage, and (c) research-oriented, practice-based comprehensive stage. The first stage took place between the 1950s and 1987 and features the initial development of the four-level model of evaluation (Kirkpatrick 1959a, 1959b, 1960a, 1960b). In this period there was an unconscious attempt at developing techniques for this little-understood topic of evaluation. Much confusion abounded among practitioners and academics about what needed to be done and even the original author seemed to be "unclear about the role that the model would play" (Wang and Spitzer 2005, p. 6).

The second stage of process-driven operational activity took place against the backdrop of globalization and international competition and saw the rise of the ROI movement (Burkett 2005; Phillips 1995, 1996; Phillips and Phillips 2002). This movement was given impetus by constant pressure from management for proof of

Table 9.1 Comparison of emphasis in nomothetic and ideographic approaches

Nomothetic methods	Ideographic methods
Deduction	Induction
Explanation via analysis of causal relationships and explanation by covering-laws	Explanation of subjective meaning systems and explanation by understanding
Generation and use of quantitative data	Generation and use of qualitative data
Use of various controls, physical or statistical, so as to allow the testing of hypotheses	Commitment to research in everyday settings, allowing access to, and minimizing reactivity among, the subjects of research
Highly structured methodology	Minimum structure

business returns from training investment. HRD academics and practitioners responded by seeking to justify the expenditure in HRD with methods for calculating ROI from training initiatives.

The third stage of evaluation has moved to context and began in 1996 with a radically new approach to evaluation. In his article, *The Flawed Four-Level Evaluation Model* Holton (1996) succeeded in creating a new agenda for research and for practitioners. Holton introduced a concerted effort to move the discussion away from outcomes as had been the case with the Kirkpatrick and Phillips' models in the preceding years. Holton suggested that by introducing context in the form of research into the transfer system, it would be possible to develop evaluation methods that were grounded in theory and also of practical value to the practitioner (Holton 1996). Other major contributions to the research on transfer have subsequently been advanced (Kontoghiorghes 2001, 2002, 2004; Tracey and Tews 2005). These three stages in the development of evaluation are worthy of discussion and development.

9.2.2 Philosophical Approaches to Evaluation

Most of the evaluation research has been conducted using a highly positivist and result-driven approach. It is conventional to position the various approaches to research along a continuum of increasing rigor. At one end is laboratory-type experimentation, and at the other, field research. The former is often known as the scientific method or positivism and draws upon structured methods copied from the natural sciences. At the other end of the continuum is the inductive tradition that uses ethnographic methods. This approach rejects the positivist tradition in favor of methods that help give richer insights in areas where subjective meaning and context play a major role. In between these poles are numerous methods that have been used by researchers to combine elements of the two traditions.

These two traditions have also been called nomothetic and ideographic. Nomothetic methods base research on systematic protocol and technique and use methods employed in the natural sciences. Ideographic methods analyze the subjective accounts derived from deep involvement in the research situation. Following Gill (1996), Table 9.1 compares the main points of nomothetic and ideographic methods.

Debate on the method of evaluation has gone through a number of phases in recent times. This progression can be shown as a kind of continuum from scientific to phenomenological approaches. While the scientific approach concerns itself with using the scientific method—being objective, quantitative, looking for scientific proof, using measures, controls and statistics, being rigorous—phenomenology concerns itself with individuals' perceptions of reality and the meaning which people attribute to their experiences. Easterby-Smith (1986) identifies three phases of this progression as scientific, systems, and naturalistic approaches.

Phenomenological evaluation tends to concentrate on how individuals perceive their experience. It is totally context specific and cannot be generalized to other individuals or to a community at large. Most approaches to the evaluation of training, in both the economics and HRD literatures, have been positivist in nature, attempting to establish causation between the independent variable (training) and the dependent variable (some organizational good or outcome). However, because there are so many intervening variables between the training and the outcome, positivist approaches have limited diagnostic utility for the human resource practitioner. It is difficult, therefore, to identify the source of problems if outcomes are not favorable. An approach is needed that will specify the intervening variables and their effects and establish a means for their measurement.

9.2.3 Kirkpatrick's Four-Level Model

The most popular and most enduring contribution to the field of evaluation has proven to be the model developed by Kirkpatrick in a series of four articles for the American Society for Training & Development Journal (Kirkpatrick 1959b, 1960a, 1960b). In these articles, Kirkpatrick outlined his four-step model of *reaction, learning, behavior*, *and results*. Perhaps because of its simplicity and ease of understanding it has become the most widely known and accepted approach to the subject among practitioners (Alliger and Janak 1989; Bates 2004; Salas and Cannon-Bowers 2001). Such has been the influence of this model that, many years later, Kirkpatrick could claim, with considerable justification, that very little had changed, in terms of content, since 1959 (Kirkpatrick 1994).

Kirkpatrick's step one was termed as *reaction* and is commonly measured soon after a training program. This step refers to the way trainees "like" and "feel toward" a program of training. Although this measurement is often referred to derisively by trainers as smile sheets or happy sheets, this practice attempts to measure the participant's reaction to the program.

Step two measures *learning* or the amount the participants believe they have learned. Kirkpatrick defines this step as measuring principles, facts, and techniques understood and absorbed by the trainees.

Step three is termed *behavior* and refers to the behavior change that has happened since the training and is defined as using learned principles and techniques back on the job.

Step four is *results* and this simply refers to a measurable impact of the training on the performance of the organization and is referred to by Kirkpatrick as results desired, including reduction of costs, reduction of turnover and absenteeism, reduction of grievances, increase in quality and quantity of production or improved morale.

The popularity of the model has been phenomenal and can be explained by several factors. Firstly, it has provided a language for talking about evaluation of training and has given practitioners a simple-to-understand systematic model for undertaking evaluation (Shelton and Alliger 1993). Secondly, it introduced a connection between the work of HRD professionals and the results of the business through its encouragement of the development of techniques to measure the impact of the training's results. If the training function is to become a true business partner it must begin to demonstrate where it is contributing to the overall results of the organization (Bates 2004). Lastly, Kirkpatrick's model simplifies (and perhaps oversimplifies) for practitioners what is complex.

The emphasis on outcomes de-emphasizes the contextual nature of a learning event which is nested within a system such as is a modern organization. The countless variables which affect human and organizational performance are not addressed in the model and thus the four-level-model of evaluation appears to have a simple and seductive appeal to the busy practitioner.

In his early articles Kirkpatrick used the term *steps* to describe the four elements of his model (Kirkpatrick 1959a, 1959b, 1960a, 1960b). Subsequently in the literature this model became known as the four-level approach to training evaluation. In these early stages Kirkpatrick was probably unconscious of the major effect that his model would have on the world of HRD. In the early years he may not have intended it to be more than a "heuristic for training evaluation" (Alliger and Janak 1989). Wang suggests that these early stages of development of the field of evaluation were atheoretical, pointing out that Kirkpatrick was confused about the role that his scheme would play (Wang and Spitzer 2005). An examination of Kirkpatrick's early articles suggests that there is some merit in Wang's assertions.

However, in more recent times Kirkpatrick asserted the implied causal linkages in the model from step to step thus:

> if training is going to be effective, then it is important that trainees react favourably. (Kirkpatrick 1994, p. 27)

> without learning, no change in behaviour will occur. (Kirkpatrick 1994, p. 51)

Kirkpatrick thus, alters his conceptualization of the model from taxonomy to a theory of training evaluation. Kirkpatrick's model has achieved a dominant position in the HRD marketplace and has achieved widespread and enduring popularity (Alliger and Janak 1989). The field of industrial and organizational psychology has adopted this model in great measure (Cascio 1987), and Kirkpatrick has popularized the training evaluation concept and created a convenient language for facilitating communication in evaluation. This popularity and dominance of the field may be due, in part, to the simplicity of the model. Practitioners find it easy to understand

Training	
Leads to	Reactions
that lead to	Learning
that leads to	changes in behavior
that lead to	changes in the organization
that lead to	change in the achievement of ultimate goals

Fig. 9.1 Bramley's cause and effect linkages. (Bramley 1991)

and yet at the same time there may be some misunderstandings, over-generalizations, and invalid assumptions (Alliger et al. 1997).

Over the period of time, since the development of the model, certain implicit assumptions within it have become more explicit in the literature. It is now common to see what Kirkpatrick termed as *steps* now being described as *levels* (Goldstein 1986). This implies that there is now a perceived integration between elements heretofore seen as independent. Since this terminology is now pervasive in the literature on evaluation, this author will use the term *levels* from this point forward. Given that these implicit assumptions exist and have been given voice, it perhaps is useful to examine to what degree they can be supported by evidence from the literature.

The first assumption is that there are causal linkages in the model. Bramley (1991) asserts that a cause and effect chain links the levels specified in these approaches (Fig. 9.1):

For pragmatic reasons it may be necessary for the training department to provide training that trainees like (otherwise trainees will not be inclined to attend for training unless forced to). However, this does not demonstrate that liking leads to learning. In fact, it may be the case that only when trainees experience challenge to the point of discomfort do they learn (Alliger and Janak 1989).

In general, it seems plausible that reactions have a relationship with the other levels of the model. There may also be some merit in positing relationships between the other levels. Learning achieved on a training course should relate to behavior since some knowledge of the subject may be a prerequisite to transfer. Similarly behavior transfer should have a relationship with results since some action is required to create an impact on the organization's metrics.

A second assumption in the literature is that the fourth level is the most significant (Aragón-Sanchéz et al. 2003; Kirkpatrick 1994). Training is an investment and companies will be interested to find out if the return from training has exceeded the investment cost (Bee and Bee 1997; Cascio 1987). This assumption has a plausibility about it that seems almost beyond question. Yet, it also seems likely that some training initiatives may not lend themselves comfortably to level four of Kirkpatrick's model. Training which is aimed at morale building or simply as an energizer may have outcomes which are either intangible or which do not sit easily in the fourth level of the model.

9.2.4 Amendments and Developments to Kirkpatrick's Model

Many evaluation models have been submitted to the literature since the 1950s. Almost without exception, each one builds on the four-level model. Where these authors differ from Kirkpatrick is in dividing the fourth level into two distinct elements thus proposing a fifth level.

Hamblin suggests a fourth level termed *organization* and a fifth level termed *ultimate value* (Hamblin 1974). Organization refers to immediate issues such as productivity or quality improvement. Ultimate value refers to profitability, survival, or growth.

AQ5

Brinkerhoff adds two levels to the four levels by including formative evaluation of the training needs and training design (Brinkerhoff 1989). Kaufman and Keller (1994) also propose a five-level model. However, in this case the fifth level is the benefits to society delivered by the training.

Phillips (1995) too contributes a model with five levels. In this model the fourth level indicates the results achieved by the organization such as productivity or quality improvement and the fifth level is ROI from the training. Cascio (1999) provides a model that differs in essence from the four levels by emphasizing performance change with a dollar value estimation of that performance change.

Kirkpatrick (1994, p. 54) was still able to state that "*content has remained basically the same.*" It is difficult to argue with this assertion. Bramley (1991) also notes that the evaluation of training remains dominated by the four-level approach of reaction, learning, behavior, and results.

9.2.5 Contributions of the Kirkpatrick's Model

Kirkpatrick's four-level model has popularized the training evaluation concept (Wang et al. 2002). Its principal contribution is that it has focused attention on the issue of outcomes from training interventions (Broad and Newstrom 1992). It has also shown that single outcome measures cannot reflect the complexity of training interventions and has emphasized the importance of using multiple measures of training effectiveness (Bates and Holton 2004). The model indicates the aspects and outcomes one should examine and assess when evaluating training programs (Wang and Spitzer 2005).

Today, increasing emphasis is placed on evaluating training outcomes, and the four-level model offers the practitioner community a vocabulary for discussing the variety of training outcomes that can actually be measured. The model also offers practitioners some sophistication for assessing training interventions, especially where organizations are used to making assessments in simplistic, reaction-based terms. Furthermore, practitioners are introduced to the notion that their training programs actually do affect the strategy of the organization, offering them central and powerful roles that might be denied to them were they to be perceived merely as a support function organizing training events.

For the academic community, the Kirkpatrick's model gives a point of reference for future research. This model, in its early days, epitomized the atheoretical stage in the history of evaluation (Wang and Spitzer 2005). From its atheoretical limitations, however, many academics have found their points of departure into rich fields of research (Holton 1996; Tracey and Tews 2005).

9.2.6 *Criticisms of the Kirkpatrick's Model*

Although Kirkpatrick's model is dominant, its application is less than complete. In one study, some authors noted that evaluation practices have changed very little in the last 30 years for which data is available (Twitchell et al. 2000). Few companies calculate the ROI from employee training in an effective and reliable manner. Bartel (2000), in a review of the literature on ROI research, found that a lack of data and poor methodology rendered conclusions difficult.

Critics of the four-level model have attacked it for perceived flaws which include its incompleteness and the failure to establish causal linkages (Bates 2004). These criticisms are now examined.

The Kirkpatrick's model may be termed incomplete in terms of its application and scope. Firstly, it is not universally applied by practitioners. An American Society for Training and Development (ASTD) study found that 77% of the organizations surveyed used reaction measures, 38% evaluated learning, 14% measured behavior transfer, and only 7% carried out evaluations at the level of results (Van Buren and Erskine 2002). Either organizations believe that reaction measures are the most powerful (a debatable proposition) or they do not have the ability and/or the will to invest the time and effort into evaluating other criteria. Secondly, because it concentrates on outcomes, the model tends to ignore elements that gave rise to and surround the training program. Thus, there is a risk that any failure to achieve outcomes may be attributed to the intervention itself (Holton and Naquin 2005).

The term *reaction* is also used in the original model to describe a single construct (Kirkpatrick 1959a). However, it has been demonstrated that there are two elements to reaction: *affective reaction* and *utility reaction*. Affective reaction refers to liking the training, whereas utility reaction refers to perceived value of the training in helping them to do their job (Alliger et al. 1997).

There are also serious questions to be answered, such as the absence of essential elements from the model. The major intervening variables that affect learning such as trainee readiness, motivation, training design, and reinforcement of training on the job are not specified in the four levels (Holton 1996). In addition, individual differences may also affect outcomes and these are not specified in the model.

Kirkpatrick's model commenced its life as a taxonomy. In the early stages the author seemed to view it merely as a set of separate and unlinked steps to good practice in the evaluation of training programs. However, he later claimed that there were causal linkages in the model (Kirkpatrick 1994). This assertion has not been supported by the literature (Alliger and Janak 1989).

In general, reactions, either *affective* or *utility*, do not correlate with learning (Alliger and Janak 1989; Dixon 1996). Some argue that reactions should not be regarded as a primary outcome but, rather, as a moderator of the relationship between training motivation and learning (Mathieu et al. 1992). This is in direct opposition to the four-level model where trainee reactions, defined as happiness, are a primary outcome of training (Kirkpatrick 1994).

It has been argued that the four-level approach is no more than a taxonomy of outcomes (Holton 1996). This reflects (Alliger and Janak 1989; Alliger et al. 1997) who, in two comprehensive studies stated that the implied causal linkages between each level of taxonomy had not been demonstrated by research. Their literature reviews show that reported correlations between Kirkpatrick's levels have varied widely. They noted, however, that utility reaction measures related more strongly to learning and performance transfer than affective measures (Alliger et al. 1997). Counter-intuitively, they also suggested that *utility* measures are more predictive of transfer than learning measures.

Most research into relationships between the levels of the four-level model has indicated weak connections between the reaction level and other levels (Alliger 1989; Alliger et al. 1997; Dixon 1996). However, Warr et al. (1999) suggest that such conclusions are not appropriate for links between reactions and learning when more differentiated indicators of reaction are examined. Four measures of trainee reactions were taken and were found to be associated consistently with measures of learning (Warr et al. 1999).

Donald Kirkpatrick's typology was and remains the dominant framework for listing training criteria for evaluation. However, there have been criticisms and questions regarding its effectiveness as an evaluation approach (Kaufman and Keller 1994; Holton 1996). The current practice and theory of evaluation do not answer sufficiently well the questions that trainers and others have about organizations' training and development efforts (Preskill 1997).

AQ8

Research into the four-level model suggests that it does not comprise the elements required to describe it as a theory. For example, various meta analyses and other research have found virtually no relationship between trainee reactions and the other levels (Dixon 1996; Alliger 1989; Alliger et al. 1997). Such studies fail to establish the direct relationship often implied by Kirkpatrick and his followers between the levels of the model, the most common being the assumption that reactions can be used as a surrogate measure for training effectiveness. However, as Tannenbaum and Yukl (1992, p. 425) suggest: "liking does not imply learning."

This model generally also fails to take account of the dynamic nature of training and development, or the important conditions that await the trainee in the workplace on his/her return from the training intervention. Kirkpatrick's approach cannot account for the reasons for choosing the intervention and the process of nomination of the trainee for that intervention. This model does not ascertain if the training process has taken place in an atmosphere conducive to the development of the right attitudes on the part of the learner. It does not ask if the learner, on returning to the workplace, will be given the required level of support and be given the opportunities to test out the new knowledge in a supportive atmosphere.

9.2.7 Conclusions and Future Research

Evaluation of training today has its roots in the United States where over the last century educators began to use learning objectives as tools of evaluating their work. Authors have noted the different stages in the development of evaluation, (a) the atheoretical stage, (b) the process-driven operations stage, and (c) the research-oriented, practice-based comprehensive stage. During the past 50 years the development of evaluation has reflected a wider debate in the social sciences in terms of epistemology from interpretivist to positivist approaches. The dominance of one particular model in the practitioner field has led to controversy.

Despite its longevity, the evaluation profession does not have a set of effective and widely used tools for practitioners and researchers (Bates 2004). It is also disturbing that a 50-year-old model, under constant attack by the academic profession and many leading practitioners, is still being promoted by the largest practitioner organization, the ASTD (Paradise 2007).

However, it can also be said that over the past 50 years the measurement and evaluation of HRD has come of age. Today, it can be described as an issue of major importance in HRD, a "topic of debate" (Phillips and Phillips 2002). The debate seems to sustain itself with continued momentum. Even today, researchers find value in durable model of the four levels of evaluation (Smidt et al. 2009). Although there does not seem to be any flagging of interest in the issue, it is less certain that HRD researchers and practitioners are clear about the direction of evaluation.

HRD needs research and new directions on evaluation criteria. The Kirkpatrick model needs to be replaced by an alternative, grounded in research but of practical use for the practitioners.

Research into its replacement has commenced and is described by Wang and Spitzer (2005) as the research-oriented, practice-based comprehensive stage. This stage heralds the introduction of context by several authors in search of approaches to supersede the Kirkpatrick model (Holton 1996; Tracey and Tews 2005; Kontoghiorghes 2004). A new vocabulary has been developed including the arrival of such terms as the *transfer climate* and *transfer system* incorporating a range of factors that help and hinder the transfer of learning from training interventions back into the workplace.

Further research needs to be conducted into the factors that affect transfer of learning. Current research has been mostly situated in America and further research in the North European situation is needed (Van der Klink et al. 2001). Research till date has also used participants' self-reports as the main estimation of transfer. More concrete measures of the effective transfer of training are required. Furthermore, transfer research till date has neglected the role of the trainer as a factor in enhancing transfer of learning. It is likely that this has a significant bearing on the effective transfer of the training.

In the next section, a key development of the Kirkpatrick model is discussed—ROI. This model of evaluation attempts to place a value on the outcomes of training as a percentage return on investment figure. It gives a focus and direction to those

who seek to demonstrate financial value to the firm of HRD. It also attracts criticism from those who believe that there are too many variables involved to isolate one particular effect of training interventions.

9.3 Measuring Return on Investment

9.3.1 Introduction

Measuring ROI from training interventions has become one of the most challenging and intriguing issues facing the HRD and performance-improvement field (Phillips 2005). It is a topic which appears on many HRD conference and convention agendas and articles appear regularly in HRD practitioner and research journals dealing with the issue. Yet, there is more to be done as others note:

> While significant improvements have been made in the evaluation of training … more work could be done at the results level. (Olsen 1998, p. 74)

In meeting this need many books and articles have also been written on the subject and many consulting firms offer services to clients in the area of calculating ROI.

The issues that are driving this increased interest are emanating from the business arena. Pressure is being exerted by clients and senior management to show results from training investment (Rowden 2005). Competitive economic pressures also are causing scrutiny of expenditures, including all training and development costs. It is already clear that organizations are "shaving every expense that does not promise a return" (Ruona et al. 2002, p. 218). Systemic initiatives such as total quality management, business re-engineering, and Six Sigma have created a renewed interest in measurement and evaluation including that of training interventions. A general trend toward accountability for all staff groups is causing some HRD departments to measure their contribution. These and other factors have created a movement toward applications of an ROI process. HRD professionals must better demonstrate bottom line impact (Swanson 2000).

AQ9

9.3.2 Research on Return on Investment

ROI is one of the most intriguing issues HRD is facing today (Subramanian et al. 2012). Much of the research into ROI in training interventions has been led by ASTD. In 1994 ASTD began to collect and publish case studies in ROI. This initiative has become such a success with the practitioner community that it is now the Society's largest seller among all of its publications. The interest reflects Society's own view that the number one global trend facing HRD practitioners is developing the ROI in training (Van Buren and Erskine 2002).

Research studies continue to show the growth of interest in ROI (Matalonga and Feliu 2012). In a survey of 35 members of the International Federation of Training and Development Organizations (IFTDO), measuring ROI was consistently rated as the topic of greatest importance among members of these organizations (Phillips 1999). Perhaps the most comprehensive study in this subject in recent years was conducted by the US Corporate Leadership Council involving 278 organizations (Drimmer 2002). This study showed that 78 % of organizations saw ROI as desirable, rating it as either important or very important as a desired metric. However, only 11 % of them were using ROI as a measure of training effectiveness. These results were the same for development interventions (nontraining interventions).

Another major study attempted to determine how organizations measure the impact of corporate universities (Phillips 2000). This was a detailed benchmarking study to examine how major corporate universities are dealing with the accountability issue and ROI. It concluded that best practice sites were moving toward utilizing various techniques of evaluation including ROI. It was also concluded that these corporate universities were struggling with the problem of how to calculate ROI and what to do with the results.

One of the most visible signs of the advancement of ROI is the development of the ASTD ROI Network. Founded in 1996 by a group of practitioners, its purpose is to promote the science and practice of individual and organizational measurement and accountability. Membership is global and in 1992 it was acquired by ASTD who now operates it as an internal division. Its services are open to all members as an ASTD membership option.

The number of conferences is often a useful indicator of trends, and a variety of conference providers have concentrated on the topic of ROI in recent times. These include the International Quality and Productivity Center (IQPC) who routinely offer conferences on ROI, sometimes five per annum across the globe. ASTD ROI Network has now conducted nine annual conferences on this topic. Since 2002, ASTD has introduced the practice of having a special conference on ROI within its own International Conference and Exposition. The American Productivity and Quality Center (APQC), and the Institute for Industrial Relations (IIR) have also offered conferences in the US, Canada and Europe on ROI.

9.3.3 The Phillips' Model of ROI

The most widely known of the approaches to ROI in HRD is the Phillips' method of ROI, developed by Phillips 30 years ago. Phillips' ROI model is positivist in its approach and has gained popularity among managers.

It has been suggested that this model is an extension of the Kirkpatrick model but this has been contested in the literature as being a misconception (Wang and Wang 2005). This model has become widely accepted in the practitioner community and its strengths include the way it attempts to isolate the effects of the program from

Table 9.2 Definitions of evaluation levels in the Phillips' model

Sr. No.	Level	Brief description
1	Reaction and planned action	Participants react to the program and make plans to transfer the learning
2	Learning	This assesses changes in skills, knowledge, or attitude change
3	Application and implementation	Measures back on the job behavior change
4	Business impact	Measures tangible changes in the business as a result of the program
5	ROI	Calculates the ROI of the program including costs and benefits

other influences. The evaluation levels used in the model are broadly analogous to the steps in Kirkpatrick's taxonomy (Kirkpatrick 1994). However, there is an additional level of ROI in the Phillips's model. The definitions of the levels of Phillips' approach are shown in Table 9.2.

Level one measures the reaction of the participants to the program as does the Kirkpatrick taxonomy and others, but this model includes an action plan for implementation of changes in work practices based on the learning achieved in the program. Level two is identical to other outcomes-based evaluation models in that it purports to measure the knowledge, skills, and attitudes that have been acquired on the program. These may be tests, role plays etc. Level three, action and implementation, uses a variety of follow-up methods to determine whether participants applied on the job what they have learned. Level four is business impact, and the measurement here focuses on the metrics which the program itself was designed to change. Typical level four measures include output, quality and costs etc. Level five is described aspirationally as the "ultimate evaluation" (McArdle 2011, p. 249). This measure compares the monetary benefits of the program with the program costs.

Phillips demonstrates how to place monetary values on training's worth and calculate the ROI of a training intervention. Phillips' approach is to collect post program data, and then to isolate the effect of training from other influences and thereby attempt to estimate, in financial terms, the contribution made by the training intervention. The sequence of this method is as follows:

- Develop a baseline of performance
- Conduct the program
- Collect postprogram data
- Isolate the effects of the program
- Convert benefits to monetary value
- Calculate the ROI

9.3.4 Evaluation Planning

In the ROI model there are three specific elements of planning which are important to the success of the application of the model (purpose, feasibility, and objectives). These elements are outlined in this section.

Purpose Several distinct purposes can be identified in evaluation of HRD interventions (Phillips 2003, p. 37).

- Improve the quality of the learning and outcomes
- Determine whether a program is accomplishing its objectives
- Identify the strengths and weaknesses in the learning process
- Determine the benefits/cost analysis of an HRD program
- Assist in marketing HRD programs in the future
- Determine whether the program was appropriate for the target audience
- Establish a database, which can assist in making decisions about the programs
- Establish priorities for funding

Purposes may often determine the scope of the evaluation so these should be identified in advance of the development of the evaluation plan. When practitioners are planning an ROI evaluation, for example, the purposes include perhaps comparing the costs and benefits of the program. This purpose has significant implications for the type of data collected, the data collection methods, and the means of communicating the results.

Feasibility When planning the ROI impact study, it is necessary to decide upon the appropriate levels for evaluation. An evaluation project may stop at level three where all that is required is a report on the extent to which the staff actually uses what they have learned. Other studies need to go to level four where the consequences of staff behavior in terms of the impact on the metrics of the organization are considered. This level four study will seek to find both hard and soft measures linked to the program. In the end, if an ROI calculation is needed, then the impacts on the metrics of the organization must be converted to monetary data so that an ROI formula can be used and a percentage figure obtained. For the ROI study to be achieved, a feasibility study is usually carried out. Typical questions at this stage of assessing feasibility are as follows (Phillips 2003):

- What specific measures have been influenced with this program?
- Are those measures readily available?
- Can the effect of the program on those measures be isolated?
- Are the costs of the program readily available?
- Will it be practical, and feasible, to discuss costs?
- Can the impact data be converted to monetary value?
- Is the actual ROI needed or necessary

These questions are important to help the evaluation team decide what is possible and appropriate in terms of the levels of evaluation that can be accessed in the project.

Table 9.3 Sample data collection form. (Adapted from Phillips and Phillips 2001)

Level	Program objectives	Data collection method	Data sources	Timing	Who is responsible
1	Reaction, Satisfaction and Planned Actions Positive reaction–four out of five	Questionnaire	Trainee	End of program	Trainer
2	Learning Learn to use communication skills with customers	Observation of practice in class	Trainer	During class	Trainer
3	Application and Implementation Initial use of five simple skills 80% of trainees use all skills with all customers	Follow-up session Follow-up questionnaire	Participant Participant	3 weeks after program Three months later	Trainer Line manager
4	Business Impact Sales increase	Business data figures	Company records	Three months after end of program	Line manager
5	ROI 30%	A figure of 30% ROI gives management some comfort that ROI is planned for.			

Objectives As seen in Table 9.3, programs are evaluated at different levels. The level of evaluation achieved corresponds to the level of the objectives set for the program

- Reaction, affective, and utility objectives
- Learning objectives for knowledge, skill, and attitudes
- Application and behavior objectives
- Impact objectives
- ROI objectives

Every evaluation exercise requires that objectives be identified prior to the execution of the program. Learning objectives are traditionally developed for training programs but other levels such as application and impact levels are not, however, necessary they may be for the calculation of ROI and evaluation of results.

Objectives of the program are deeply connected to the front end-training needs analysis of the program. After the business need is determined, the analysis determines the performance that is required to deliver on this need. Different objective types link directly to a different but appropriate level of evaluation. If the application and impact objectives are not available, then they must be developed.

The next part of the planning stage of the Phillips' model is the use of planning documents (data collection plan, ROI analysis plan, and the project plan) and these are discussed next.

Data Collection Plan A data collection plan is a document for the recording of the major elements and issues in respect to the collection of data for the four evaluation levels. An example of such a plan is shown in Table 9.3 and is drawn from an evaluation project in sales training (Phillips and Phillips 2001).

In this planning document broad areas for planning are acceptable. At a later point, more specific objectives will be developed. In the measures column the specific measure is listed and in the method column the actual technique used to collect the data is also listed. The origin of the data is listed in the source column and the timing indicates the scheduling of collection. The responsibilities column indicates who will collect the data.

ROI Analysis Plan This document captures information on items that are needed to develop the ROI calculation. Table 9.4 shows a completed ROI analysis plan for the sales program which was discussed in Table 9.3.

In the first column in Table 9.4 is listed the critical data which will be used to calculate the ROI. In the second column, the method used to isolate the effects of the training in the calculation of ROI is listed next to each of the data items in the first column. The conversion column tells how the information will be converted to monetary values so that the calculation for ROI can be made. The cost categories are listed in the fourth column. Normally these will be consistent across all training courses; however, in certain circumstances, there may be cost items which are specific to a particular course and these will be noted here. In the fifth column, intangible benefits are listed which are expected from the program and this list can be generated through discussions with the various stakeholders. The targets for communications are listed in the sixth column. Out of the many targets that could be listed, Phillips lists four that are "always recommended."

- Top management group
- Line manager of trainees
- Trainees themselves
- Training and development staff

These groups are typical stakeholder groups who need to know about the results of an ROI analysis. In the final column other elements which might influence the program implementation or which might be crucial to note in the conduct of the ROI analysis are noted. Typical among these might be the degree of access to sources of data, unique analysis issues such as contact with control groups and ability issues concerning participants (Phillips 2003).

Project Plan The third planning document necessary for the ROI initiative is the project plan. This document is generic in the sense that most executives who are required to execute an organizational project would be familiar with and utilize a project plan. It comprises a description of the program, its duration, target audience, and number of participants. The timeline of the initiative will be shown also from the inception to the final communication of ROI results to the stakeholders listed earlier. A project plan is a common tool to control any given project. The critical element of time usually drives a project plan. If senior management has a specific

Table 9.4 Sample ROI analysis plan. (Phillips 2003, p. 44)

Data Items	Methods of isolating effects of the program	Methods of converting data	Cost categories	Intangible benefits	Communications targets	Other influences and issues
Weekly sales per associate	Control group analysis	Direct conversion using profit contribution	Facilitation fees	Customer satisfaction	Program participation	Job coverage during training
	Participant's estimates		Program materials	Employee satisfaction	Electronics department managers at targeted stores	Communication with control group
			Meals and refreshments		Senior store executives district, region, headquarters	Seasonal fluctuations
			Facilities		Training staff: instructors, coordinators, designers, and managers	
			Participant's salaries and benefits			
			Cost of coordination			
			Evaluation			

583 end date in mind, then this will be agreed and consequently all other dates in the
584 project plan are fixed in respect of this conclusion of the project. For this purpose,
585 a generic project planning tool will suffice.
586 The planning documents described above (the data collection plan, the ROI anal-
587 ysis plan, and the project plan) can be used as a basis for the direction of the ROI
588 study. The documents enable the key decisions required during the planning phase

to be made. Subsequent to this, is the execution of the project but this is merely a formulaic implementation of the decisions made in the earlier phase of the initiative.

Two types of data are collected in applying the ROI methodology: hard and soft. Hard data comprise output, quality, cost, and time measures. Soft data comprise job and customer satisfaction. A variety of methods are used to collect including these:

- Questionnaires and surveys
- Simple tests
- Observation of performance on the job
- Interviews with trainees
- Focus groups
- Performance data

The collection of data will be constrained by issues such as time and budget. Nonetheless, care should be taken to select the method appropriate to the specific program and the setting.

9.3.5 Isolating the Effects of the Training

One of the difficulties in evaluating training interventions is determining or attributing causality. Given that there are so many variables which have an impact on organizational metrics, any evaluation attempt must respond to the challenging possibility that alternative explanations exist for the improved performance other than the training one. As a result, with any method it is important to address this issue, especially one such as ROI that deals with impacts on the organization which occur long after the training intervention has taken place.

The objective of this stage of the model is to determine the amount of improvement following the training that is directly related to the program itself. If this can be achieved, then the calculation of ROI becomes a more precise and accurate exercise. There are many techniques, familiar to the experienced researcher, which are utilized to address this issue.

- Control group
- Trend lines
- Forecasting model
- Participant estimate
- Supervisors of participants estimate
- Senior management estimate
- Subject matter experts

These tools may be used as a comprehensive set of techniques to answer the challenge of isolating the effect of the training on the performance metrics.

9.3.6 Converting Data to Monetary Values

An ROI calculation requires that the data collected at the impact level (level 4) is converted to a monetary figure and then compared to program costs. Thus, when the impact on the results, which is attributed to training, is established it must then be translated into monetary amounts which can then be used in the ROI formula. There are many techniques available to convert these data depending on various training situations. The principal techniques amongst these are converting the profit contribution or the cost savings into monetary value. This reminds us of the critical nature of planning, of how the training is designed to affect the business performance. If this is not decided in advance then it is difficult to see how the ROI can be calculated with confidence.

Because of its importance this step is vital. However, the size of the challenge can be underestimated especially where soft data is concerned. Using a multiple approach with these strategies can increase confidence levels in the results.

9.3.7 Tabulating Costs of the Program

Tabulating costs of the program involves first of all gaining agreement in respect of the costs to be tabulated. Once this is established, this part of the model involves monitoring or developing all of the costs of the program targeted for ROI calculation. Some sample items which may be included in a cost calculation are as follows:

- Cost of designing the program
- The costs of program materials
- Trainer costs including preparation and delivery time
- Cost of facilities, rooms, technology etc.
- Travel and subsistence costs for attendees and trainees
- Salaries and overhead charges of participants

9.3.8 Calculating the Return on Investment

The formula for ROI calculation in the Phillips' method of ROI is executed using the program benefits and costs as shown below:

$$\text{ROI} = \frac{\text{Net Program Benefits}}{\text{Program Costs}} = \frac{(\text{Benefits} - \text{Costs})}{\text{Program Costs}} \times 100. \quad (1)$$

ROI is traditionally reported in many investment situations as earnings divided by investment. ROI, as a percentage, will vary according to the specific type of program being considered. Sales, supervisory, and management training can have a high ROI (frequently over 100 %) while the same calculation for technical and operator training can be lower (Phillips 2003).

9.3.9 *Criticisms of the Phillips' Model*

The Phillips' method of calculating ROI which was developed during what was described as the "atheoretical phase" of the development of evaluation approaches, and subsequently formed its centerpiece, has been described as a "noteworthy milestone" (Wang and Spitzer 2005, p. 7). Many practitioners regard this ROI technique as the ultimate goal of evaluation and an addition of a fifth level to the Kirkpatrick model of evaluation.

This stage focused almost entirely on the operational processes of evaluation. This method enabled HRD professionals to derive and obtain a percentage figure reflecting the impact of the HRD intervention on their workplace.

The Phillips' method of ROI has unfortunately been associated with the Kirkpatrick model in a misconceptualization by researchers and practitioners alike. The labeling of this model as the "fifth level of evaluation" deems it to be as extension of the Kirkpatrick model, and with support from its author, has been termed the "ultimate level of evaluation" (Phillips 2003, p. 12). This does not add light to the evaluation landscape.

The Kirkpatrick model does not contain any specific techniques or step by step approaches to conduct the evaluations at each level. The implied causal linkages between the levels do not stand up to scrutiny (Alliger and Janak 1989). Thus, the Kirkpatrick model is not really a theoretical model but rather a taxonomy (Holton 1996). ROI analysis, on the other hand, by itself is a technique to measure the financial returns for HRD interventions. It is conceptually inappropriate to link the Kirkpatrick model and the Phillips' technique and adds further confusion among HRD professionals.

The rise to prominence of this method has, however, through an extensive emphasis on ROI, been significant in terms of increasing the awareness of both functional management and HRD practitioners about the importance of evaluation for HRD interventions, emphasizing the importance of HRD investment in organizations and motivating further efforts in the pursuit of credible evaluation approaches (Wang and Spitzer 2005).

The Phillips model can be criticized for an over emphasis on financial data as many training interventions are aimed at developing intangible outcomes (Wang et al. 2002). Some suggest that more qualitative factors should be given more weight (Burke and Hutchins 2007). Others bemoan the "moment in time" aspect of the ROI calculation which tends to ignore the time factor in the development of ROI.

Although ROI has been used to calculate the returns from various forms of training including software training the cost factors are usually known as the organization is usually collecting this data already (Diaz and Sligo 1997). Conversely, benefits are much more difficult to identify and there needs to be a considerable level of agreement around the accepted assumptions as to what benefits entail and which are acceptable to quantify. The Phillips model does not offer much guidance as to how this agreement may be reached. Some authors also criticize Phillips for the use of an average figure for ROI suggesting that a more subtle approach is required such as the use of statistical process control tools to measure the variation before and after the intervention (Matalonga and Feliu 2012).

9.4 Conclusions and Future Research

There is mounting evidence that the work environment and training climate has had an increasing relevance and this has moved attention toward broader and more integrative models of training evaluation which involve the transfer of learning (Holton 1996; Tracey and Tews 2005). These authors propose an alternative model involving the context surrounding the training intervention. Here, they avoid the weaknesses that were identified in the case of outcomes-based models that assumed simple relationships and causal linkages were in place.

Other writers have invoked expectancy theory to develop models of transfer that move the field beyond the outcomes-based approaches of the Kirkpatrick model; however, not all the factors which affect transfer have been identified (Kontoghiorghes 2004).

Organizations, therefore, are anxious to demonstrate that the investment in HRD is delivering reasonable returns and methods for demonstrating this value have been considered for many years. Critics have suggested that outcomes-based methods of evaluation have failed to deliver both theoretically and practically for organizations. Some authors have suggested that the transfer system may offer potential for development. Till date there has been a significant degree of research into transfer of learning but some factors remain elusive. Research is needed to identify factors heretofore unrecognized and to identify the relative importance of these factors and to further ascertain to what degree context plays a role and to what degree the importance of transfer factors alters with context.

The ROI method is used as a surrogate for transfer of training from the training intervention back into the workplace. The method can provide an objective and consistent measure of the effectiveness of HRD interventions across different training programs and different business sectors. What is now required, is a means of identifying and measuring the factors which affect transfer of learning so that variations in these factors could be compared to variations in the impact of training. This topic could move the debate concerning evaluation from "does training work?" to a question of "how training works?"

References

Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology, 42*(2), 331–340.

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

Aragón-Sanchéz, A., Barba-Aragón, I., & Sanz-Valle, R. (2003). Effects of training on business results. *Journal of Human Resource Management, 14*(6), 956–980.

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

Bartel, A. P. (2000). Measuring the employer's return on investments in training: Evidence from the literature. *Industrial Relations, 39*(3), 502–524.

Bates, R. (2004). A critical analysis of evaluation practice: The Kirkpatrick model and the principle of beneficience. *Evaluation and Program Planning, 27*(3), 341–347.

Bates, R., & Holton, E. F. III (2004). Linking workplace literacy skills and transfer system perceptions. *Human Resource Development Quarterly, 15*(2), 153–170.

Bee, F., & Bee. R. (1997). *Training needs analysis and evaluation*. London: Institute of Personnel and Development.

Bramley, P. (1991). *Evaluating training effectiveness: Translating theory into practice*. London: McGraw.

Brinkerhoff, R. O. (1989). *Achieving results from training*. San Francisco: Jossey-Bass.

Broad, M. L., & Newstrom, J. W. (1992). *Transfer of training*. Reading: Addison-Wesley.

Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human Resource Development Review, 6*(3), 263–296.

Burkett, H. (2005). ROI on a shoestring: Evaluation strategies for resource-constrained environments or ROI on a shoestring. *Industrial & Commercial Training, 37*(2), 97–105.

Cascio, W. F. (1987). *Applied psychology in personnel management*. Englewood Cliffs: Prentice-Hall

Diaz, M., & Sligo, J. (1997). How software process improvement helped Motorola. *Software, IEEE, 14*(5), 75–81.

Dixon, N. M. (1996). New routes to evaluation. *Training and Development, 50*(5), 82.

Drimmer, A. (2002). *Reframing the measurement debate: Moving beyond program analysis on the learning function*. Washington, DC: Corporate Executive Board.

Easterby-Smith, M. (1986). *Evaluating management development, training and education*. Aldershot: Gower.

Gill, J. J. P. (1996). *Research methods for managers*. London: Paul Chapman.

Goldstein, I. L. (1986). Training in organisation: Needs assessment, development and evaluation. Monterey: Brooks/Cole.

Hamblin, A. C. (1974). *Evaluation and control of training*. New York: McGraw-Hill.

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

Holton, E. F. III., & Naquin, S. (2005). A critical analysis of HRD evaluation models from a decision-making perspective. *Human Resource Development Quarterly, 16*(2), 257–280.

Kaufman, R., & Keller, J. M. (1994). Levels of evaluation: Beyond Kirkpatrick. *Human Resource Development Quarterly, 5*(4), 371–380.

Kirkpatrick, D. L. (1959a). Techniques for evaluating training programs. *Journal of ASTD, 13*(11), 3–9.

Kirkpatrick, D. L. (1959b). Techniques for evaluating training programs: Part 2—Learning. *Journal of ASTD, 13*(12), 21–26.

Kirkpatrick, D. L. (1960a). Techniques for evaluating training programs: Part 3—Behaviour. *Journal of ASTD, 14*(1), 13–18.

Kirkpatrick, D. L. (1960b). Techniques for evaluating training programs: Part 4—Results. *Journal of ASTD, 14*(2), 28–32.

Kirkpatrick, D. L. (1994). *Evaluating training programs*. San Francisco: Berrett-Koehler.
Kontoghiorghes, C. (2001). Factors affecting training effectiveness in the context of the introduction of new technology—A US case study, 248–260
Kontoghiorghes, C. (2002). Predicting motivation to learn and motivation to transfer learning back to the job in a service organization: A new systemic model for training effectiveness. *Performance Improvement Quarterly, 15*(3), 114–119.
Kontoghiorghes, C. (2004). Reconceptualizing the learning transfer conceptual framework: Empirical validation of a new systemic model. *International Journal of Training and Development, 8*(3), 210–221.
Matalonga, S., & Feliu, T. S. (2012). Calculating return on investment of training using process variation. *IET Software, 6*(2), 140–147.
Mathieu, J., Tannenbaum, S., & Salas, E. (1992). Influences of individual and situational characteristics on measures of training effectiveness. *Academy of Management Journal, 35*(4), 828–847.
Olsen, J. H., Jr. (1998). The evaluation and enhancement of training transfer. *International Journal of Training and Development, 2*(1), 61–75.
Paradise, A. (2007). *State of the industry: ASTD's annual review of trends in workplace learning and performance*. Alexandria: ASTD.
Phillips, J. J. (1995). Return on investment-Beyond the four levels. In E. Holton III (Ed.), *Academy of HRD 1995 conference proceedings*. Baton Rouge: Academy of HRD
Phillips, J. J. (1996). *Accountability in human resource management*. Houston: Gulf.
Phillips, J. J. (1999). HRD trends worldwide: Shared solutions to compete in a global economy. Houston: Gulf.
Phillips, J. J. (2000). *The corporate university: Measuring the impact of learning*. Houston: American Productivity and Quality Center.
Phillips, J. J. (2003). *Return on investment in training and performance improvement programs*. Burlington: Butterworth-Heinemann.
Phillips, J. J. (2005). Communicating results to top executives. *Chief Learning Officer, 4*(4), 60–68.
Phillips, P. P., & Phillips, J. J. (2001). *Action: Measuring Return on Investment*. Alexandria: American Society for Training and Development.
Phillips, J. J., & Phillips, P. P. (2002). Technology's return on investment. *Advances in Human Resource Development, 4*(4), 512–532.
Preskill, H. (1997). HRD evaluation as the catalyst for organizational learning. In E. Holton III (Ed.), *Proceedings of the Academy of Human Resource Development*. Baton Rouge: AHRD.
Rowden, R. W. (2005). Exploring methods to evaluate the return-on-investment from training. *Business Forum, 27*(1), 31–36.
Ruona, W., Leimbach, M., Holton, E. III., & Bates, R. (2002). The relationship between learner utility reactions and predicted learning transfer among trainees. *International Journal of Training and Development, 6*(4), 218–228.
Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology, 52*(1), 471–499.
Shelton, S., & Alliger, G. M. (1993). Who's afraid of level 4 evaluation? A practical approach. *Training and Development Journal, 47*(6), 43–46.
Smidt, A., Balandin, S., Sigafoos, J., & Reed, V. A. (2009). The Kirkpatrick model: A useful tool for evaluating training outcomes. *Journal of Intellectual & Developmental Disability, 34*(3), 266–274.
Subramanian, K. S., Sinha, V., & Gupta, P. D. (2012). A study on return on investment of training programme in a government enterprise in India. *Vikalpa: The Journal for Decision Makers, 37*(1), 31–48.
Tannenbaum, S. I., & Yukl, G. (1992). Training and development in work organizations. *Annual Reveiw of Psychology, 43*(3), 399–441.
Tracey, J. B., & Tews, M. J. (2005). Construct validity of a general training climate scale. *Organizational Research Methods, 8*(4), 353–374.

Twitchell, S., Holton, E. F. III., & . Trott Jr., J. R. (2000). Technical training evaluation practices in the United States. *Performance Improvement Quarterly, 13*(3), 84–110.

Tyler, R. W. (1949). *Basic principles of curriculum design*. Chicago: University of Chicago Press.

Van Buren, M. E., & Erskine, W. (2002). *The 2002 state of the Industry report*. Alexandria: American Society for Training and Development.

Van der Klink, M., Gielen, E., & Nauta, C. (2001). Supervisory support as a major condition to enhance transfer. *International Journal of Training and Development, 5*(1), 52–63.

Wang, G. G., & Spitzer, D. R. (2005). Human resource development measurement and evaluation: Looking back and moving forward. *Advances in Developing Human Resources, 7*(1), 5–15.

Wang, G. G., & Wang, J. (2005). Human resource development evaluation: Emerging market, barriers, and theory building. *Advances in Developing Human Resources, 7*(1), 22–36.

Wang, G. G., Dou, Z., & Li, N. (2002). A systems approach to measuring return on investment for HRD interventions. *Human Resource Development Quarterly, 13*(2), 203–224.

Warr, P., Allan, C., & Birdi, K. (1999). Predicting three levels of training outcome. *Journal of Occupational and Organizational Psychology, 72*(3), 351–375.

Chapter 9: Author Query

AQ1. Please check if the following edited sentence “The transfer of training was, as a research area, originally focused…” retains its intended meanin.

AQ2. “Fig. 9.1” is not cited in the text. Please check whether we have inserted the citation in the right pla e.

AQ3. “Aragon, Antonio, Barba-Aragon, & Sanz-Valle 2003” has been changed to “Aragón-Sanchéz et al. 2003” to match the reference list. Please confi m.

AQ4. “Bee 1997” has been changed to “Bee and Bee 1997” to match the reference list. Please con irm.

AQ5. “Kaufman et al. 1994” has been changed to “Kaufman and Keller 1994” to match the reference list. Please confi m.

AQ6. “Twitchell and Holton 2000” has been changed to “Twitchell et al. 2000” to match the reference list. Please confi m.

AQ7. “Van Buren 2002” has been changed to “Van Buren and Erskine 2002” to match the reference list. Please confi m.

AQ8. “Preskill and Holton 1997” has been changed to “Preskill 1997” to match the reference list. Please confi m.

AQ9. “Swanson 20 0 a d McArdle 20 1” re cited in the text but is not given in the reference list. Please provide full references or delete the citations.

AQ10. “Tyler 1950” has been changed to “Tyler 1949”. Please check and con m