Evaluating human resource interventions

Author:
Joha Louw-Potgieter[1]

Affiliation:
[1]Section of Organisational Psychology, University of Cape Town, South Africa

Correspondence to:
Joha Louw-Potgieter

Email:
joha.louw-potgieter@uct.ac.za

Postal address:
Section of Organisational Psychology, University of Cape Town, Rondebosch 7701, South Africa

Orientation: Programme evaluation is a transdiscipline, which examines whether a programme has merit or not. A programme is a coherent set of activities aimed at bringing about a change in people or their circumstances.

Research purpose: The purpose of this special edition is to introduce readers to the evaluation of human resource (HR) programmes.

Motivation for the study: There are few comprehensive evaluations of HR programmes despite many publications on functional efficiency measures of HR (i.e. measures of cost, time, quantity, error and quality).

Research design, approach and method: This article provides a value chain for HR activities and introduces the reader to programme theory-driven evaluation.

Main findings: In summarising all of the contributions in this edition, one of the main findings was the lack of programme evaluation experience within HR functions and the difficulty this posed for the evaluators.

Practical/managerial implications: This introductory article presents answers to two simple questions: What does HR do? and, What is programme evaluation? These answers will enable practitioners to understand what programme evaluators mean when we say that programme evaluation seeks to determine the merit of a programme.

Contribution/value-add: The main contribution of this introductory article is to set the scene for the HR evaluations that follow. It alerts the reader to the rich theory contribution in HR literature and how to apply this in a theory-driven evaluation.

Introduction

In 2006, the Section of Organisational Psychology at the University of Cape Town (UCT), established a Master's degree option in programme evaluation. As is usual with the implementation of new academic programmes, this initiative was contested from within and without the section of Organisational Psychology. Those within Organisational Psychology argued that programme evaluation did not belong in its domain, whilst those outside Organisational Psychology claimed it for their domains. It seemed that programme evaluation belonged everywhere and nowhere in the University of Cape Town. This is quite common in the case of a transdisciplinary programme. According to Scriven (2003), a transdiscipline is a discipline with its own methods, which can be applied to other disciplines or domains or knowledge. Like most domain disputes, this debate is far from over. In the meantime, the Master's degree in programme evaluation has become a popular and sought-after degree for South African and international postgraduate students.

There are specific reasons for this high demand for programme evaluators in South Africa and Africa. Firstly, within the public sector in South Africa, the Department of Monitoring and Evaluation was created within the National Planning Commission in the Presidency to monitor the delivery of social programmes. In order to comply with the requirements of this department, a monitoring and evaluation function within national and regional government departments has been mooted. Secondly, within the private sector in South Africa, there is sporadic evidence of evaluation activity within service (e.g. human resource, financial services) and social responsibility departments (Field, 2011). Thirdly, in Africa, big donor organisations (e.g. the President's Emergency Programme for AIDS Relief, Centers for Disease Control, the European Union, Swedish International Development Cooperation Agency) include evaluation as part of their funded interventions. These evaluation endeavours are good news for the Universities of Cape Town and Stellenbosch, the only two universities in South Africa that offer degrees in programme evaluation.

Purpose of the special edition

In this special edition, we present the evaluations of human resource programmes produced within the Section of Organisational Psychology at UCT. This work reflects the pioneering spirit of the Section in its quest to establish programme evaluation as a standard assessment within the HR function. It is also a tribute to the Section's academics and students who ventured into the relatively unchartered waters of 'real world evaluation' (Bamberger, Rhug & Mabry, 2006) within the human resource (HR) domain.

Current theoretical perspectives

For a long time a myth persisted that the value of HR interventions could not be measured (Fitz-enz, 1995). This state of affairs has changed and measures of transactional efficiency (i.e. measures of cost, time, quantity, error and quality) abound within the HR function (Fitz-enz, 2000). Therefore, the need is not for more measures, but for systematic and consistent measures that will enable organisations to make informed decisions about the merit of HR interventions (Boudreau & Ramstad, 2007). Evaluations, or decisions about programme merit, are still the exception rather than the norm in HR. Skinner (2007) commented that:

> given that [evaluation] is so central to what people do, it is perhaps surprising that evaluation, as a planned and formal activity, appears to be so problematic in an organisational context. (p. 118)

Edwards and his colleagues (Edwards, Scott & Raju, 2003, 2007) have documented a number of HR programme evaluations, but otherwise this area of evaluation is still developing.

Firstly, one could speculate that the reason for the paucity of evaluation of HR programmes is that most non-HR programme evaluators perceive this function as consisting of multiple, unrelated people management tasks. This perception is reinforced by HR's generation of many independent transactional efficiency measures. Secondly, the reason could be that most HR and non-HR people have not made the connection between the purpose of an evaluation and the specific HR intervention. Is the purpose of the evaluation to find out how the intervention was implemented, or is it to determine the results that it produced? Can we use the information from a systematic evaluation to make an informed decision that established whether or not to improve or discontinue the programme?

The main aims of this introductory article are:

- to simplify the activities of the HR function into an explanatory framework that enables non-HR people to understand what HR does
- to demystify the domain of programme evaluation into five comprehensible steps
- to show how HR interventions can be evaluated.

How does human resource work?

When confronted with the myriad of people management tasks that HR practitioners engage in daily, most people fail to find logic in all of this detail. However, there is a simple but powerful logic (sometimes also called a value chain) that categorises these tasks and presents the categories in a sequential fashion.

One can think of people management practices as a sequence of activities that follows the employee throughout the organisation, namely:

- recruitment and selection (a person applies for a job, goes through a selection process, is placed in a suitable job and inducted into the organisation)
- pay, benefits and reward (depending on the job level, the person is paid a specific salary with benefits like medical aid and a pension fund and may be recognised by means of non-financial awards like employee of the month)
- training and development (the person is trained for the specific job and developed for personal growth or organisational change)
- performance management (after a suitable period, the person's performance in the current job is reviewed and evaluated against appropriate organisational standards, a process which is often linked to pay and a training or development plan)
- employee relations (by means of fair practices, health and safety programmes and interventions such as diversity programmes, the organisation strives to foster a workplace with a positive culture and climate for its employees).

This value chain is presented in Figure 1.

A number of authors have attempted to categorise HR tasks and make sense of the function. In this regard, the work of Fitz-enz (2000) stands out. He presented six typical management activities of the HR function, namely: planning, acquiring, maintaining, developing, retaining and evaluating. According to Fitz-enz, evaluating is not a separate HR function, but integral to the effective functioning of the other five HR activities. He then continues to describe what happens in each of these activities:

- Planning provides guidelines to human resource needs and succession planning as intended at a particular time.
- Acquiring refers to recruitment and selection, both from within and outside the organisation.
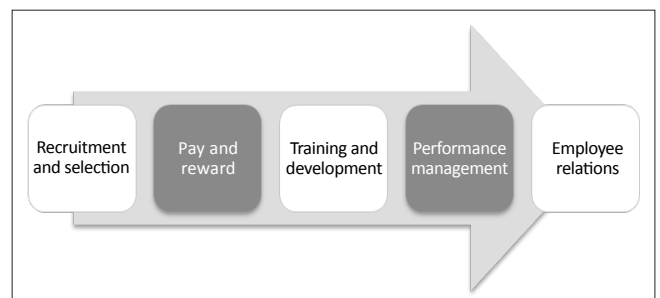- Maintaining refers to how employees are maintained by means of reward and recognition.

FIGURE 1: Human resource value chain.

- Developing focuses mainly on how training and development programmes are used to develop people to their fullest potential.
- Retaining, the last activity, involves programmes for staff retention, which may focus on employee relations, organisational culture and values.

These two models show significant overlap and clarify the purpose of HR. They deal with superordinate categories of work, which simplifies HR work for non-HR people. At the same time, they enable programme evaluators to see at a glance how HR activities are interlinked and add value to employees' journeys within the organisation.

Do human resource departments work?

This is a typical evaluation question that assesses the value or merit of HR interventions. In order to get to HR evaluation we must first explore a model that will enable us to judge the merit of HR activities. This model shows a hierarchy of steps, which can be used in evaluating HR programmes; it is also possible to isolate some of the steps and only deal with those in an evaluation.

The model was developed by Rossi, Lipsey and Freeman (2004) and is presented in a step-wise hierarchy in Figure 2.

This model is usually applied to social programmes (i.e. programmes that seek to improve social conditions like poverty, hunger, crime, etc.). However, it can also be applied to any programme that consists of an organised set of activities, which seeks to change the current state of affairs. For instance, the model could be used to assess a programme that aims to bring about improved performance in the workplace, and that consists of organised skills training, performance management and reward components.

The model starts with the first step, need, which is the motivation of most social programmes. Why do we want to implement a programme? Because there is a problem. What exactly is the problem? How big is it? How is it distributed? An example of a people management need can illustrate this. A company is experiencing significant staff turnover. The problem is that employees, especially managers, are leaving the company. This is a serious problem, as the company has lost 15% of its middle and senior managers in the last six months. Ten per cent of this 15% manager loss is in the Johannesburg branch. There is a need to do something about this problem. A programme to stop managerial job loss could be a solution to this problem.
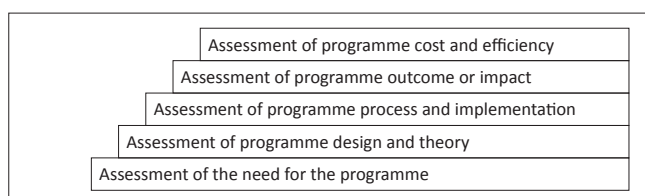
Source: Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation. A systematic approach.* (7th edn.). Thousand Oaks: Sage

FIGURE 2: Step-wise model of programme evaluation.

The second step in the hierarchy is programme theory and design. Programme theory refers to the assumptions or ideas of stakeholders about how the programme works (Bickman, 1987), or how it will bring about change. Very often, programme theories are modest theories (Donaldson, 2007). These programme theories may also be implausible. It helps considerably to have experts who know the subject area to assist in building a more complex or a plausible programme theory. The better practitioners understand how a programme will bring about change, the better they can choose relevant programme activities that will support this change. In this way, we can design a programme, which shows good alignment between programme theory and programme activities. An example of a modest programme theory is when stakeholders assume that training will cause improved organisational performance. If this theory is left unchallenged, a skills training programme may be implemented and when evaluated, may show little impact on company performance. However, if some research was undertaken on the relationship between training and performance, the programme managers might have realised that training is a necessary but not a sufficient intervention for improved performance (Brinkerhoff, 1998). They may also have come to the conclusion that, in order for training to be effective, employees first need to apply what they have learned in training to their jobs. Furthermore, they may have read that employees who have supportive supervisors (Noe, 2005), who encourage experimentation, may create a department where people try out their new skills without fear of being laughed at or punished for mistakes made. Thus, these programme managers, armed with a more plausible theory about the link between training and performance, may now set about to design a programme with training, application and supervisory support components.

The third step in the model is implementation. Implementation addresses coverage (who received the programme), process (how the programme was implemented) and programme resources (whether or not there were enough resources to implement the programme properly) (Rossi *et al.*, 2004). Many newcomers to programme evaluation find these implementation terms confusing, mainly because there is more than one label for the same concept. For instance, an implementation evaluation is often referred to as a process evaluation, coverage is also called service utilisation, process and service delivery are used interchangeably, and resources are often called support and organisational functions. In this special edition, we shall use the terms implementation evaluation with the sub-categories of coverage, process and programme resources.

Many well-designed programmes with robust programme theories fail, because they are poorly implemented. An example of a well-designed programme with poor implementation is the South African national housing programme. In the HR function, implementation is even more problematic. Often a programme is conceptualised by senior HR staff, but implemented by line managers who have

not been part of the design process or who have not been trained to execute the implementation (Purcell & Hutchinson, 2007). Should an HR programme evaluation conclude that the implementation was done poorly, this can invariably be related to unclear accountability for its implementation.

The fourth step deals with the outcomes and impact. Outcomes refer to a change in the state of the problem, a change in the state of affairs, or a change within the recipient of the programme (Rossi *et al.*, 2004). Outcomes can be short, medium or long-term. Impact refers to a causal relationship and addresses the question of whether or not the programme, and not anything else, has brought about the change in the problem, state of affairs or the recipient. We often expect significant impact from HR programmes. However, researchers like Cohen (1988) have indicated that a realistic expectation for the effect of a well-targeted behaviour change programme is approximately a 30% – 35% change or improvement.

The fifth and final step deals with programme cost. Please note that this is not about budgeting for a programme. Budgeting happens when people plan programmes. This step deals with judging how much a specific programme costs per recipient, or how much a specific programme costs in comparison with other programmes of the same kind (Rossi *et al.*, 2004). Sometimes a programme might bring about the desired change, but may simply be too costly to sustain. High-level management or executive development programmes often fall within this category.

Monitoring and evaluation are often used simultaneously in book titles or university programmes. Monitoring refers to tracking the implementation, outcomes or both, of a programme over time. When programme evaluators monitor outcomes, we do so by means of indicators (the best representations of the outcome) (Kusek & Rist, 2004). For instance, we may monitor the training outcome, knowledge acquisition, by means of performance on a knowledge test. Sometimes we add standards to our outcome indicators. A standard is a measure which tells us whether or not our outcome progress is good enough (Kusek & Rist, 2004). In HR, we have a number of standards, which indicate whether or not performance is good enough. These standards are often called benchmarks and organisations like Saratoga or companies undertaking salary surveys update and publish these regularly. Sometimes these standards form part of international agreements, like the International Labour Organisation's conventions on hours of work, maternity protection, or minimum age. At a national level, these standards are reflected in the Basic Conditions of Employment Act. Or we can set our own realistic standards. For instance, what percentage of absenteeism is acceptable? Does this percentage hold for Friday and Monday absenteeism too (when more staff are absent), or do we inflate the standard for these days?

Again, the terminology used for monitoring might be confusing for newcomers to programme evaluation. Often, indicators are called criteria; some programme evaluators distinguish between indicators and measures whilst others assume that an indicator includes a measureable component; often, standards are called targets or yardsticks.

Sometimes an outcome cannot be measured directly, thus, we use proxy indicators. A proxy indicator is an indirect representation of the outcome (Kusek & Rist, 2004). In South Africa, we often refer to race to indicate socio-economic class or previous disadvantage. Another example of a proxy indicator is the following: in the 2009 elections, a survey included 'number of bathrooms' as a proxy indicator of socio-economic class. Socio-economic class was then used to predict for which political party people within that class category would vote.

When we are dealing with social programmes, we may use pre-designed indicators (Kusek & Rist, 2004), like the Millennium Developmental Goals or the International Monetary Fund's Financial Soundness Indicators. In HR, pre-designed indicators are provided in salary surveys (i.e. this is the basic salary for a process engineer in Gauteng) or from HR practice (i.e. what is the optimal HR staff:employee ratio in a manufacturing organisation?).

Often, programme evaluators use baseline indicators. A baseline indicator provides information at the beginning of the monitoring period and serves as a starting point for future performance (Kusek & Rist, 2004). An example of an HR baseline indicator is a measure of skill or knowledge taken prior to the implementation of a skills or knowledge training programme. After the training, another knowledge or skill measure is taken and compared with the baseline measure to judge whether or not a trainee has acquired skill or knowledge during the training.

Usually, a social programme is aimed at improving a social condition, in other words, it does good (Rossi *et al.*, 2004). Recipients of social programmes are often referred to as beneficiaries, and are the people who experience the social benefits of the programme. These beneficiaries are now less poor, have better access to health service delivery, have houses, or live in peace.

When we evaluate HR programmes, the beneficiaries are defined less clearly. Who are the beneficiaries of an HR programme which is introduced to improve work performance, but which does not have a remuneration component tied to it? Who benefits from a selection programme aimed at providing the best person-job fit? Who benefits from a training programme in which the skills to be learned are organisation-specific and not transferrable to other companies? With HR programmes, the beneficiary is often the organisation and not the employees. Sometimes both benefit, sometimes one party benefits more than the other. Most HR programmes are not aimed at doing good; they are aimed at improving organisational performance and recipients may therefore be a more realistic term for the people who receive the programme.

Rationale of this special edition

In this special edition, we show that, by using the methods of programme evaluation, we can improve current HR programmes or determine the merit of these programmes. Unlike most texts, we do not just exhort HR practitioners to measure or evaluate; we provide clear examples of how to improve or judge the merit of HR programmes. We present seven HR programme evaluations, which cover the full HR value chain, from staffing to employee relations.

In the recruitment and selection category, an evaluation of an induction programme is described (Hendricks & Louw-Potgieter). This is a theory evaluation that showed how a modest programme theory led to the development of sparse programme activities, which did not produce the outcomes envisaged by the programme manager. A plausible programme theory and more extensive programme activities were suggested, which might result in organisational identification and staff retention.

What is still missing is an evaluation of a selection programme in the recruitment and selection category. Some organisations use psychological assessments for manager selection, whilst others have developed assessment centres or have committed to a series of focused interviews. We do not know which of these selection programmes are most effective. Also, an organisation seldom uses all three types of selection programme simultaneously. This makes comparison difficult. Furthermore, the absence of a plausible programme theory of manager selection, which includes most of the relevant mediating and moderating variables, complicates useful evaluation.

In the pay and reward category, Salie and Schlechter showed how to work around unexpected obstacles in a formative evaluation of a recognition programme. These authors presented an elegant exposition of the difference between the programme manager's and the recipients' perceptions of this programme. This evaluation highlighted the importance of asking both the programme designer and recipients whether or not a programme works.

The crux of the reward and recognition category is whether or not pay motivates employees and which kinds of pay are the best motivators. This is a complex area to evaluate and would require in-depth knowledge of types of pay and employee motivation. For the present, we leave this as a challenge for a prospective doctoral student.

In the training and development category, we found the highest number of evaluations (three). It seems that training evaluation is the default in HR evaluation (in South Africa, and also elsewhere in the world). This in itself is interesting, as training is not a high-cost HR intervention like recruitment and selection or pay. So, why would most HR evaluations focus on training? Perhaps the pervasive popularity of Kirkpatrick's (1994) four levels of evaluation could explain this. Kirkpatrick offered the HR function

a four-category evaluation model (reaction to training, learning, application of training and effect of training) which appealed intuitively to all. Today, reaction to training is used as a measure for virtually every training programme offered within organisations. Most HR professionals proudly offer information regarding the number of delegates and whether these delegates liked the training or not, as 'evaluations' of training. Sometimes learning is assessed, but application and the effect of training are seldom evaluated. The assessment of the latter two levels is more complex and cannot be achieved by means of simple questionnaires. For this reason, training evaluation seems stuck with the number of delegates and their reaction to the training.

Firstly, the initial evaluation in the training and development category (Buys & Louw) highlights how a plausible programme theory might be misaligned to the initial needs assessment done for a supervisor development programme. Whilst the programme under evaluation could be judged as successful, it did not fulfil the initial need of reducing the cost of supervisor recruitment. Rather, it had become a training programme for incumbent supervisors.

Secondly, Rundare and Goodman show how additional programme activities like group learning enhanced the effectiveness of a perinatal care programme for midwives. In their recommendations for future evaluations of the programme, they indicated how a quasi-experimental design could strengthen future evaluations.

Thirdly, Beets and Goodman used Brinkerhoff's Success Case Method (SCM) (2003, 2006) to evaluate whether or not recipients had applied the skills, knowledge and values (SKAs) they acquired on a coaching training programme in their work. The evaluators illustrated here how to use this method and made useful suggestions on when not to use it.

In the performance management category, Joseph, Emmett and Louw-Potgieter use an implementation evaluation to show how a pay-for-performance programme had little effect because of its flawed implementation. They extracted the essential variables for successful implementation of such a programme. This evaluation has been used by the organisation in question to re-launch the pay-for-performance programme during 2009.

Finally, in the employee relations category, Duffy and Louw present a plausible programme theory for a wellness intervention. Like Salie and Schlechter, these authors also indicated how to work around evaluation problems when programme staff change their commitment to evaluation. Duffy and Louw intended to do an implementation evaluation, but had to settle for a theory evaluation when the economic downturn prevented the organisation from rolling out the wellness initiative.

The evaluations presented in this special edition utilised a theory-driven evaluation approach (Donaldson, 2007). A theory-driven evaluation approach requires a good

description of the programme activities. From these activities, 'a plausible and sensible model of how a programme is supposed to work' is constructed (Bickman, 1987, p. 5). This model is called a programme theory and it is usually tested against existing social science research in order to ascertain whether or not it constitutes a plausible theory of change (i.e. how the programme will change the recipients or the problem). Apart from detailing how a theory-driven approach to evaluation science works, the articles in this special edition also show how to craft evaluation reports for different audiences, how to undertake evaluations for different purposes, how to utilise different evaluation methods and how to overcome the challenges of conducting evaluations of HR programmes in organisations. In utilising this approach, we hope to contribute to improved HR practice in South Africa.

Acknowledgements

Competing interests

References

Bamberger, M., Rugh, J., & Mabry, L. (2006). *RealWorld evaluation: Working under budget, time, data, and political constraints*. Thousand Oaks: Sage.

Bickman, L. (1987). The functions of programme theory. *New Directions for Program Evaluation, 33*, 5–18. http://dx.doi.org/10.1002/ev.1443

Boudreau, J.W., & Ramstad, P.M. (2007). *Beyond HR. The new science of human capital.* Boston: Harvard Business School Press.

Brinkerhoff, R.O. (1998). Clarifying and identifying impact evaluation. In S.M. Brown & C.J. Seider (Eds.), *Evaluating corporate training. Models and issues,* (pp. 141–166). Boston: Kluwer Academic Publishers. http://dx.doi.org/10.1007/978-94-011-4850-4_7

Brinkerhoff, R.O. (2003). *The success case method. Find out quickly what's working and what's not*. San Francisco: Berrett-Koehler Publishers.

Brinkerhoff, R.O. (2006). Getting real about evaluation. *Training & Development, 60*(5), 24–25.

Cohen, J.W. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates.

Donaldson, S.I. (2007). *Program theory-driven evaluation science*. New York: Lawrence Erlbaum Associates.

Edwards, J.E., Scott, J.C., & Raju, N.S. (2003). *The human resources program-evaluation handbook*. Thousand Oaks: Sage.

Edwards, J.E., Scott, J.C., & Raju, N.S. (2007*). Evaluating human resources programs: A six-phase approach for optimizing performance.* San Francisco: John Wiley & Sons.

Field, C. (2011). *Training evaluation practices in the South African financial services sector.* Unpublished manuscript. Section of Organisational Psychology, University of Cape Town.

Fitz-enz, J. (1995). *How to measure human resources management*. New York: McGraw-Hill Inc.

Fitz-enz, J. (2000). *The ROI of human capital. Measuring the economic value of employee performance*. New York: Amacom.

Kirkpatrick, D.L. (1994). *Evaluating training programs: The four levels.* San Franciso: Berret-Koehler.

Kusek, J.Z., & Rist, R.C. (2004). *Ten steps to a results-based monitoring and evaluation system.* Washington, DC: The World Bank. http://dx.doi.org/10.1596/0-8213-5823-5

Noe, R.A. (2005). *Employee training and development*. Boston: McGraw-Hill.

Purcell, J., & Hutchinson, S. (2007). Front-line managers as agents in the HRM-performance causal chain: Theory, analysis and evidence. *Human Resource Management Journal, 17*, 3–20. http://dx.doi.org/10.1111/j.1748-8583.2007.00022.x

Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation. A systematic approach.* (7th edn.). Thousand Oaks: Sage.

Scriven, M. (2003). Evaluation in the new millennium: The transdisciplinary vision. In S.I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium,* (pp.19–42). Mahwah: Lawrence Erlbaum Associates.

Skinner, D. (2007). Evaluating SHRM: Why bother and does it really happen in practice? In M. Millmore, P. Lewis, M. Saunders, A. Thornhill & T. Morrow (Eds.), *Strategic human resource management. Contemporary issues,* (pp. 117–149). Harlow: Prentice Hall.