

Focus

EVALUATION IN DEVELOPMENT EDUCATION: CROSSING BORDERS

In this article, **Gilbert Storrs** seeks to identify the key elements of performance measurement systems initially adopted by the business management sector and now prevalent in the public and non-profit sector. He will examine the body of research and learning in this area and open up the possibilities of developing a system that is appropriate for application in the development education environment.

Introduction

“To throw up our hands and say we cannot measure performance in the social sectors the way you can in business is simply a lack of discipline” (Collins, 2005:7).

The continued funding of development education (DE) may in part be dependent on the development education community designing and embracing evaluation and measurement systems that can clearly demonstrate the value added impact of DE programmes. Donors, taxpayers, customers and stakeholders all demand effective and efficient use of resources that are allocated to projects and programmes, and almost all require evaluation of those projects to demonstrate effectiveness. Evaluation can also assist organisations and individuals in maintaining purpose and clarity around their mission, goals and objectives and to sustain them in the delivery of their desired outcomes.

Measurement systems offer us a feedback mechanism on what works and what needs change or adjustment. They provide valuable information on the impacts, intended and unintended, of our actions and initiatives, and as such are a vital part of all learning, living systems. Well-constructed measurement systems can ensure a consistent stream of direct and concise feedback in areas vital to the aims and objectives of a project and its success.

This article will outline the core essentials of performance management systems and place them firmly in the framework of a strategic planning process. The article will review performance systems and in particular look at the ‘balanced scorecard’ (Kaplan & Norton, 1992) as a model worthy of analysis and reflection. Rather than proposing an off-the-shelf solution the article argues

for a model that will embrace the strategic planning process to capture the added value of development education in all its aspects.

In examining the current practice of project evaluation, the article will emphasise the learning opportunities lost in contracting evaluation to external evaluators and argues that evaluation must be an integrated, on-going participatory process of measurement, reflection, adjustment and learning. Input from external evaluators, if employed at all, needs to be redirected to the project planning and design phase. The external input would focus on facilitating, through dialogue with the stakeholders (particularly the co-learners), the identification of the desired outcomes and the selection of a set of appropriate measures. The measurement and evaluation system should be designed to give continuous feedback in key areas to ensure corrective action can be taken throughout the duration of the project or organisation. Skilled internal, continuous evaluation by co-learners who are keen to learn and maximise their contribution, combined with evaluation from the perspective of the relevant stakeholders, is a rational way forward for evaluation. Evaluation is of little use if it is not planned and executed as part of an on-going, integrated and embedded process within the context of a strategic plan that contains an inspiring vision, a clear mission, and explicit measures to capture the impacts and outcomes.

Resistance to performance measurement

Performance measurement is almost universally resisted. This can arise from a fear by those being measured that they will be harshly judged, or that the measurement system adopted will not reflect the full range of contributions and added value of individuals or programmes. Some may believe that it is not possible to measure the particular type of work performed. Evaluation and measurement are often viewed negatively with mental images of judgement, competition, failure, and being monitored and observed. A more positive mental model of performance management and measurement needs to be constructed, if it is to serve us well.

Evaluation with appropriate measurement systems is a powerful asset; and the learning from that evaluation must be handed back to those who can gain most from the process:

“There is also the issue of what use will be made of the measurement. Is it intended for control and oversight, or is it for learning? Is it for the

sake of a third party or for the players involved? The useful aspect of measurement is that it helps us make explicit our intentions...‘What measurement would have meaning to me?’ This opens the discussion on the meaning of the activity and the use of the measure we take. It keeps measurement from being a supervisory device, and turns it into a strategy to support learning” (Block, 2003:22-23).

Performance appraisal

Prior to the development of strategic planning and project evaluation in the business sector, individual performance appraisal was practiced in a large number of organisations. An examination of the issues and insights surrounding the practice of performance appraisal is worthwhile when moving to consider strategic planning and project evaluation.

Perhaps the most groundbreaking piece of research into individual performance appraisal was carried out in General Electric during the early part of the 1960s. This research arose from the perception that performance appraisal was resisted and where practiced it produced little if any positive outcome.

“Surveys generally show that people think that the idea of performance appraisal is good...in actual practice it is the extremely rare operating manager who will employ such a programme on his own initiative...few indeed can cite examples of constructive action taken - or significant improvement achieved - which stem from suggestions received in a performance appraisal interview with their boss” (Meyer, et al., 1965:123).

This interesting insight may also apply to the prevailing model of project evaluation where evaluation is external to the day-to-day business and carried out at project end with little true participation by the key learners. Evaluation is often regarded as a necessary chore to be endured rather than an opportunity to learn and improve performance. Evaluations carried out under the prevailing model can reflect more the experience and perspective of the external evaluator than the objective performance of the project or organisation. Evaluation needs to be continuous and integrated into the day-to-day business and most vitally must take place in the context of a clearly documented set of desired outcomes, specific targets, explicit measures and baseline readings for

these measures. General Electric went on to investigate performance appraisal systems to find out what was effective in the conduct of performance appraisal:

“At the General Electric Company we felt it was important that a truly scientific study be done to test the effectiveness of our traditional performance appraisal programs [and] we found out some very interesting things – among them the following: criticism has a negative effect on achievement of goals; praise has little effect one way or the other; performance improves most when specific goals are established; defensiveness resulting from critical appraisal produces inferior performance; coaching should be a day-to-day activity, not a once a year activity; mutual goal-setting, not criticism, improves performance; and participation by the employee in the goal-setting procedure helps produce favourable results” (Meyer, et al., 1965:124).

General Electric completed a follow-up experiment to validate the conclusions derived from the original study and as a result introduced a system called Work Planning and Review (WP&R).

“The key elements of the WP&R system as outlined are:

- There are more frequent discussions on performance;
- There are no summary judgements or ratings made;
- Salary action discussions are held separately; and
- The emphasis is on mutual goal planning and problem solving” (Meyer, et al., 1965:127).

Another important aspect of General Electric’s WP&R system was that objectives, goals and measures were valued concepts in performance measurement even at the early stages of the development of performance management systems:

“In these WP&R discussions, the manager and his subordinate do not deal with generalities. They consider specific, objectively defined work goals and establish the yardstick for measuring performance. These goals stem of course from broader departmental objectives and are defined in relation to the individual’s position in the department” (Meyer, et al., 1965:128).

This piece of research and practice in General Electric provides foundation stones for the construction of enlightened measurement and

evaluation systems to meet the needs of all stakeholders. Our challenge is to build measurement and evaluation systems based on the body of learning and research mixed with innovation and creativity. The research uncovered some interesting insights on what works in individual performance assessment and yet these basic principles are often ignored in the design of and practice of modern evaluation systems.

Enter the ‘balanced scorecard’

In the early 1990s, David Norton, president of the information technology consulting firm Nolan, Norton & Company, and Robert Kaplan, Professor of Accounting at the Harvard Business School, proposed a new performance management framework called the balanced scorecard. This system is now in widespread use both in the private and public sector worldwide.

Their paper in the January/February 1992 edition of the Harvard Business Review began with the following:

“What you measure is what you get. Senior executives understand that their organisation’s measurement system strongly affects the behaviour of managers and employees. Executives also understand that traditional financial accounting measures like return on investment and earnings per share can give misleading signals for continuous improvement and innovation”(Kaplan & Norton, 1992:71).

The arrival of the balanced scorecard model opened up the whole area of performance measurement, evaluation and strategic planning to a new set of perspectives and a set of selected balanced measures.

“The widespread adoption of some form of a balanced scorecard approach by thousands of public service organisations since its 1992 debut is a testament to the methodology’s adaptability” (Cole & Parston, 2006:35).

Essential components of the balanced scorecard

The balanced scorecard is a model of planning and evaluation that can be used to manage organisations or projects. Norton and Kaplan proposed measurement from four perspectives: customer; financial; internal process; and innovation and learning. This balanced approach avoided the distortions

caused by previous narrow measurement systems which focused on financial measures and provided individuals and organisations with a system that could measure and capture the full range and richness of their contributions towards executing their mission and realising their vision. All evaluation was considered from a number of perspectives and a variety of measures were selected to reflect the full impact and contribution of the actions and initiatives of the operation.

“Today’s systems must have the capabilities to identify, describe, monitor, and fully harness the intangible assets driving organisational success...the balanced scorecard provides a voice of strength and clarity to intangible assets, allowing organisations to benefit fully from their astronomical potential” (Niven, 2008:5).

Kaplan and Norton not only focused on outcome measures but also stressed the importance of measuring the drivers of future growth and development under the perspectives of internal process and innovation and learning:

“The measures are balanced between the outcome measures - the results from past efforts - and the measures that drive future performance. And the scorecard is balanced between objective, easily qualified outcome measures and subjective, somewhat judgemental, performance drivers of the outcome measures...Innovative companies are using the scorecard as a strategic management system to manage their strategy over the long run” (Kaplan & Norton, 1996:10).

There are three sets of performance measures normally utilised in the balanced scorecard.

1. **Input measures** Simply track inputs. Input measures, while necessary, are the least valuable as they take no account of what happens to the inputs. Are the inputs used efficiently or effectively?
2. **Output measures** Track what comes off the end of the production line.
3. **Outcome measures** Track the benefits (hopefully) and more accurately the impacts. Outcomes often measure change in circumstances, and/or behaviour, skills and attitudes and can capture the intangibles referred to by Niven (2008:5). Outcomes may take time to manifest and are normally (but not always) measured in the longer term.

The move towards outcome measures

As measurement systems have matured, more emphasis has been placed on the necessity to focus on outcome measurement. In their publication *Public Productivity through Quality and Strategic Management*, Halachmi and Bouckaert provide a provocative yet interesting distinction between outputs and outcomes:

“In the final judgement what counts is the quality of the outcome, not the process or result of a given procedure. Education and formal education for example are not the same. Formal education is results (output) while education is an outcome...Formal credentials provide only a partial picture of one’s whole education. The history of knowledge is full of examples of individuals who were able to make tremendous contributions to the development of civilisation in spite of (or because of) their lack of formal schooling” (Halachmi & Bouckaert, 1995:6).

Identifying the desired outcomes

A major element of any strategic planning process is identifying the desired outcomes for the project. The dialogue around identifying, discussing and refining the outcome measures is invaluable in building understanding and clarity around the desired outcomes. This dialogue needs to take place at the conception and planning stage and needs full inclusion of the stakeholders and the frontline participants. In the ‘Zeroing in on Outcomes’ chapter of *Unlocking Public Value*, Cole and Parston state:

“The importance of focusing on outcomes has been recognised by public service organisations around the world as they attempt to measure, manage and improve performance” (Cole & Parston, 2006:19).

In the area of DE it is important to be aware that some of the outcomes may take considerable time to emerge. Medium and long-term participative reflections on process and outcomes are an invaluable part of any evaluation and learning process.

“We need patience precisely because deeper learning often does not produce tangible evidence for considerable time. You don’t pull up the radishes to see how they are growing” (Senge, et al.,1994:45).

Attributing outcomes

There is a need to attribute the outcomes to the initiatives/actions of the project or programme. In DE evaluation it is important to be alert to the fact that outcomes may be the result of other influences rather than directly attributable to a particular DE initiative:

“The performance information should be attributable. There should be a discussion and explanation of the extent to which the accomplishments achieved can be attributed to the activities of the program - how the program in question has made a difference” (Auditor General of Canada, 1997).

Outcomes can be both intended and unintended and a survey of outcomes must also be open to identifying unintended outcomes. Organisations often ignore or filter out the unintended outcomes particularly if they have a negative impact on the stakeholders or the environment.

The balanced scorecard perspectives

The balanced scorecard model proposed planning and evaluating initiatives from a number of perspectives. The initial model proposed four specific perspectives.

The customer perspective

Perhaps the most important of the perspectives proposed by the balanced scorecard model for the non-profit organisation is the customer perspective. The concept of customer is sometimes shunned in the non-profit sector however much can be gained from entering into a dialogue around the issue. A shift to the language of customer and customer service can have an important impact on the power dynamic of the relationships in addition to enabling more meaningful outcomes.

“A clear distinction between private versus non-profit and public sector balanced scorecards is drawn as a result of placing mission at the top of the framework. Flowing from the mission is a view of the organisations’ customers, not financial stakeholders. Achieving a mission does not equate to fiscal responsibility and stewardship, instead the organisation must determine whom it aims to serve and how their requirements can be best met” (Niven, 2008:33).

Introducing the concept of customer and customer service into the development education arena is likely to be met by resistance. The dialogue around the appropriateness of the use of the term customer can produce valuable insights for reflection into the motivation, values and intentions of the players. The dialogue surrounding who development education customers are could also prove to be worthwhile.

The key issue unfolded by Norton and Kaplan was the need to consider the perspective of persons other than those providing, controlling, or managing the inputs, actions and resources. The move was to shift the focus from provider perspective to the recipient perspective in planning and measuring performance. In DE there may be a need to make a similar shift to consider additional perspectives. The perspective of co-learners or co-creators in a learning space needs to be brought to the forefront of any measurement and evaluation system.

“Many political and educational plans have failed because their authors designed them according to their own personal views of reality, never once taking into account (except as mere objects of their actions) the men-in-a-situation to whom their programme was ostensibly directed...For the truly humanistic educator and authentic revolutionary, the object of action is the reality to be transformed by them together with other people - not other men and women themselves” (Freire, 1970:75).

How do we measure the desired outcomes of the co-learners and what form of measures will they choose to use? How and when can we establish these desired outcomes and what if they conflict with the agenda of the facilitator or the funder? These are important questions worthy of dialogue within the DE community. The balanced scorecard conceptual model would place the perspective of the co-learners at the top of the agenda.

The innovation and learning perspective

One of the arguments against measurement systems in the education field is their inability to value and measure creativity and innovation. The balanced scorecard model, however, emphasises the importance of learning and innovation in determining future value in a business environment. The DE community could well consider incorporating specific objectives, targets and measures in their strategic plans to foster and capture the richness of creativity and innovation in the learning environment. Indeed creativity should have a

role in creating the measures and evaluation systems. Placing the attributes of learning and innovation within a strategic planning framework is an essential part of any balanced planning, measurement and evaluation system. While there is a view that the subtleties and qualitative nature of learning is not conducive to measurement, Robert Chambers expresses a different view:

“A question has been asked about counting the uncountable. Participatory methods have a largely unrecognised ability to generate numbers which can also be commensurable and treated like any other statistics. Through judgement, estimation and expressing values, people quantify the qualitative. The potential of these methods is overdue for recognition” (Chambers, 2009:245).

The financial perspective

In the business environment typical financial goals and measurement are concerned with profitability, growth, and shareholder value. While these attributes are not as relevant in the DE environment, it is essential to design appropriate performance measures for the financial perspective.

“For self-preservation alone, public and non-profit agencies must demonstrate the effective stewardship of what limited financial resources they have to a confused and sceptical public as well as funding bodies” (Niven, 2008:188).

In the monograph to accompany *Good to Great: Why business thinking is not the answer*, Collins further comments on the legitimacy of financial measures while distinguishing between business and the social sector:

“For a social sector organisation, however, performance must be assessed relative to mission, not financial returns. In the social sectors, the critical question is not ‘How much money do we make per dollar of invested capital?’ but ‘How effectively do we deliver on our mission and make a distinctive impact, relative to our resources’” (Collins, 2005:5).

Cole and Parston outlined their concept of the public service value model in 2006. While the model builds on the balanced scorecard model it gives valuable insights into the application of financial measures into the non-profit sector. Their model focuses on delivering the identified desirable outcomes in the most cost effective way.

“With the Public Value Model, value is defined as producing a basket of outcomes desirable to stakeholders and doing so cost-effectively” (Cole & Parston, 2006:65).

Financial measurement is often the most resisted and resented form of measurement in the non-profit sector. Financial measurement is sometimes seen as a device for delivering on a cost-cutting agenda and at best a necessary evil required by funders. A more constructive view of financial measurement would see it as an invaluable tool to allow us to maximise our desired outcomes from the finite resources available.

“By our definition, public service value is about more than simply attaining outcomes, and it is more than just reducing costs; it is about doing both in a balanced fashion and understanding the strategic trade-offs available along the way” (Cole & Parston, 2006:63).

Financial measures when used to measure cost effectiveness have an essential place in any balanced measurement system. The financial measures must be constructed to facilitate the delivery of the outcomes desired by the organisation or project. Financial measures when adopted as part of a balanced measurement system are an invaluable aid in determining the initiatives that maximise the desired outcomes.

When an organisation has identified clear desired outcomes and associated measures including baseline readings, it is in a position to move forward to develop financial measures that will optimise the attainment of desired outcomes for the available financial input.

The internal perspective

The internal perspective looks at the internal business processes of an organisation and seeks efficiency around use of resources, product cycle time and product quality. Improving learning systems, communication and networking accompanied by a set of outcome measures, all have their place in the strategic plan. Process is also vital in any learning programme, and measurement and evaluation from the perspective of process has a place in DE.

While the original balanced scorecard provided for four perspectives, different organisations have introduced additional perspectives or changed the original perspectives to reflect their particular contribution to society. Any system of evaluation requires full participatory dialogue around its design,

creation and application; only then can the potential of feedback and learning opportunity be fully harnessed.

The importance of values in evaluation

Most mission statements contain value statements, i.e. ‘the values that will guide our actions’. Without translation into day-to-day activities and interactions, value statements remain as aspirations and have little real impact. In constructing any strategic project or business plan, it is essential to work on formulating and agreeing a set of values that will inform and guide daily operation and practice. Any evaluation process needs to facilitate and measure the translation of values from aspirations to a level where their practice makes a perceptible difference to the stakeholders and in particular co-learners.

Benchmarking and best practice

Evaluation in education has sometimes relied on benchmarking and the use of best practice exemplars. However their use raises a number of issues, such as who determines what constitutes ‘best practice’? Does the term close the door on future dialogue, preventing exploration, creativity, participation and invention? The use of benchmarking and identifying best practice has serious limitations as circumstances always differ from one environment to the next. There is a need to participate and co-create unique evaluation systems for unique and diverse circumstances.

“I caution against a reliance on [benchmarks and adopting best practice]. Your strategy map should tell the story of your strategy. The objectives you choose to represent that strategy may in some cases? mirror those of other organisations, but it’s the determination of the key drivers for your particular organisation that will ultimately differentiate your from other agencies” (Niven, 2008:159).

Benchmarking and adopting best practice can also limit learning and creativity and may foster a culture of compliance rather than innovation and discovery. This may be of particular relevance in the area of development education.

“The value of another’s experience is to give us hope, not to tell us how or whether to proceed...this is not to argue against benchmarking but to express the limits of what value we can actually find in looking elsewhere for how to proceed. Most attempts to transport human system

improvements from one place to another have been profitable for those doing the transporting – the consultants – but rarely fulfilled their promise for the end user” (Block, 2002:24).

Conclusion

The business community has long since moved in its planning and evaluation processes from narrow financial measures to balanced systems that measure outcomes from the customer, process, innovation and financial perspectives. Evaluation in the business sector is embedded in the day-to-day activities and stakeholders are actively involved in the process. Within the DE community, there is a need to design and implement measurement and evaluation systems that incorporate outcome measurement from a range of perspectives similar to what has been incorporated in the business community.

There is a need to search for the intangibles, the relationships and networks formed and strengthened, the new resilience built, the flow and transfer of power, all of these intangibles and much more need to be captured by the measurement and evaluation system. We need to focus on the essential learning opportunities afforded by participative measurement and evaluation. In the DE community especially, we need to be creative, imaginative and constantly reform and refine our evaluation systems. Most of all we need to fully engage all stakeholders in dialogue to co-create effective participative evaluation systems that serve stakeholder needs and ensure delivery of the desired outcomes:

“The purpose of impact assessment [in development education] is learning and change that makes life better for [people suffering injustices]. To achieve this, we need mixed methods and pluralism. Many approaches and tools can be, and should be, used for impact assessment. Whatever they are, they must always recognise that it is those who live in poverty, those who are vulnerable, those who are marginalised, who are the best judges and the prime authorities on their lives and livelihoods and how they have been affected. We now know, as we did not two decades ago, that they have far greater analytical capabilities than we supposed. We know that ‘they can do it’. To facilitate their own empowering analysis we now have a wealth of participatory methodologies. We need to make more and better use of them. Again and again, the injunction bears repeating: ask them!’ (Chambers, 2010:246).

Many opportunities are available in the development education community for those who are prepared to engage openly in dialogue around the formation of new and appropriate planning, management and evaluation systems that are home-made, participative, focused and aligned to give effect to a clear vision and mission within a framework of agreed values.

References

Auditor General of Canada (1997) 'April-October report', Chapter 5, Appendix B, accessed 25 June 2010, available: http://www.oag-bvg.gc.ca/internet/English/att_9705ab_e_7097.html.

Block, P (2008) *Community: The Structure of Belonging*, Berrett-Koehler Publishers, San Francisco.

Block, P (2003) *The Answer to How is Yes: Acting on What Matters*, Berrett-Koehler Publishers, San Francisco.

Chambers, R (2009) 'So that the poor count more: Using participatory methods for impact evaluation' in *Journal of Development Effectiveness*, Vol. 1, 3, pp. 245-246.

Cole, M & Parston, G (2006) *Unlocking Public Value*, John Wiley and Sons, New Jersey.

Collins, J (2005) *Good to Great and the Social Sectors: Why Business Thinking is Not the Answer*, Jim Collins, Boulder.

Freire, P (1970) *Pedagogy of the Oppressed*, Penguin, London.

Halachmi, A & Bouckaert, G (1995) 'Public Productivity through Quality and Strategic Management' in *International Institute of Administrative Science Monographs*, Vol. 1, p. 6.

Kaplan, RS & Norton, DP (1992) 'The Balanced Scorecard - Measures that Drive Performance' in *Harvard Business Review*, January/February 1992, pp. 72-79.

Kaplan, RS & Norton, DP (1996) *The Balanced Scorecard: Translating Strategy into Action*, Harvard Business School Press, Boston.

Meyer, HH, Kay, E & French, JRP (1965) 'Split Roles in Performance Appraisal' in *Harvard Business Review*, January/February 1965, pp. 123-129.

Niven, PR (2008) *Balanced Scorecard: Step by Step for Government and Nonprofit Agencies*, John Wiley and Sons, New Jersey.

Senge, P, Kleiner, A, Roberts, C, Ross, R & Smith, B (1994) *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organisation*, Nicholas Brealey Publishing, London.

Gilbert Storrs is a former public servant. During his public service career he has worked as Information Systems Manager, Project Manager, Strategy Manager, Human Resources Manager and Commissioner of Valuation. Gilbert led major transformation initiatives in the Valuation Office and worked on the development of strategic plans, and measurement and evaluation Systems. He was appointed by the Minister of Finance to the Public Sector Major Project Peer Review Group and continues to serve in this capacity. He is a member of IDEA and a practitioner in the field of development education with Communitycafé.