

Occasional Papers Series No: 44

**MEASURING
SUCCESS**

Issues in
Performance Measurement

John Hailey and Mia Sorgenfrei
INTRAC Praxis Programme

INTRODUCTION

In recent years, calls for greater accountability, concerns about quality, funding constraints and the development of a contract culture have created demands for more sophisticated performance measurement strategies. Donors and governments increasingly emphasise effectiveness and sustainability. As a consequence, NGOs are under pressure to invest more in evaluating their work and measuring its impact (Lawrie 1992; Hailey and James 2003). According to Paton (2003), non-profits today operate in an environment permeated by 'expectations of measurement'. This has led to an increasing interest in how best to design and apply new performance measurement frameworks.

The search for appropriate and effective ways of measuring performance in a variety of different contexts and cultures has raised a number of questions, including:

- Is it possible to use such frameworks in the increasingly complex and changing environment facing many development agencies today?
- Is it possible for such frameworks to be both credible and cost-effective?
- What are the operational challenges in applying performance measurement systems in NGOs and other humanitarian organisations?
- What are the issues and potential trends that may affect their development and application over the next few years?

As a result, new evaluatory strategies are being introduced that account for differing contextual and cultural differences. Across the private, public, and non-profit sectors, new performance measurement frameworks are being developed that reflect their differing needs.

This paper offers an overview of the background to these developments and assesses some of the challenges facing those implementing new performance measurement systems. We conclude by suggesting that the success of such systems depends on the way they are applied, and whether they can be used in a cost-effective and credible manner. All the evidence suggests that successful evaluation and measurement strategies depend on significant investment to ensure they are of operational value. Performance measurement systems should be seen as a valuable investment rather than a burdensome cost, if they are to be effective and credible.

WHY PERFORMANCE MEASUREMENT? ACCOUNTABILITY, PERFORMANCE IMPROVEMENT OR LEARNING

The renaissance of neo-liberal politics in the 1970s and 1980s, and moves to reduce the role of the State, resulted in an emphasis on cost-efficiency of public services and concern over whether taxpayers were getting value for money. One consequence of

this was that many development agencies (whether multilaterals, bilaterals, or NGOs) came under pressure to justify their work and document effective performance. These demands for greater accountability and provision of demonstrable results have led to the introduction of increasingly rigorous performance measures and onerous evaluatory systems. The current interest in the development of appropriate and cost-effective performance measurement systems arises from the concern to assess impact, and provide assurance that assignments have been completed satisfactorily in a cost-effective and credible manner. Concurrently, concerns about the quality of development practice, duplication of resources, insufficient organisational capacity and project failure have led to a renewed interest in developing effective performance measurement systems.

There are a number of operational benefits to effective and credible performance measures. Internally, performance measures help identify strengths and weaknesses in the organisation. Where problems are uncovered, they may highlight areas where quality is threatened and contribute to the subsequent attribution of responsibility. They may also indirectly enhance productivity by ensuring more effective use of resources and organisational capacities, as well as helping managers prioritise activities and target resources more effectively. Another advantage of such systems is that they facilitate internal communication by summarising complex situations and reducing the scope for ambiguity and misunderstanding. They provide a common language to articulate issues that need to be addressed and changes that should be made (Czarnecki 1999). Externally, such systems help identify issues of concern which may affect the organisation's legitimacy, reputation, and public profile. They may also provide evidence as to whether an organisation is fulfilling its goals and meeting stakeholder expectations. Performance measurement allows greater comparability over time and between units; it helps to identify best practice and facilitate new learning (Paton 2003).

The two principal functions of performance measurement systems are, firstly, to ensure that organisations are held accountable for their performance and actions; and, secondly, to facilitate learning and improve performance. These two roles are often at odds with each other, and tensions are provoked around this 'accountability-learning dichotomy'.

Being accountable to key stakeholders is not exclusively a non-profit and public sector phenomenon. In the private sector, firms are accountable to their shareholders, as well as customers, trade unions and other key stakeholders. Corporate governance is a matter of some debate, but it is accepted that companies have to conform to certain ethical and environmental standards, as well as legal norms. The frameworks and systems commonly used to assess the success of donor-funded development projects are based on an underlying assumption that NGOs should be accountable to their key stakeholders, most importantly to their donors and beneficiaries (Cutt and Murray 2000). But accountability is not just about donor control. It includes the fulfilment of public expectations and organisational goals, as well as responsiveness to the concerns of a wider constituency.

NGOs must seek to negotiate accountability and develop strategic approaches leading to outcomes that reflect win-win situations for the different stakeholders. In practice, they create a variety of formal and informal mechanisms to sustain stakeholder relationships, while dealing with the ongoing accountability demands (Ospina, Diaz and O'Sullivan 2002). Negotiated accountability provides NGOs and stakeholders (all of whom have their own agendas) with some insight into how well an organisation is doing, and allows them to verify whether it is achieving its goals

and fulfilling its purpose. This in turn opens up the organisation's performance to public scrutiny, and so ensures a degree of transparency.

Organisational learning and innovation are recognised as critical to organisational success and sustainability. The use of evaluatory tools and frameworks as a basis for learning and change belies the picture that they are merely instruments of managerial control or mechanisms for public accountability. For example, such systems as benchmarking are used as much to exchange ideas about performance improvement and innovation. Unfortunately, in the development sector, the donor imposition of performance measurement on NGOs is all too commonly seen as a means to placate the demands of auditors and ensure a level of accountability rather than learning and performance improvement. The donors would deny this and argue that they too are deeply concerned with shared learning and dissemination of good practice. There is, however, a widespread belief that donors use such measures to fulfil formal requirements rather than initiate organisational learning (CDRA 2001).

This tension is increasingly well understood, and strategies such as Results Based Management are being introduced to integrate performance feedback into future strategy and promote organisational learning. However, while many development agencies have experience in developing performance measurement frameworks, little documentation exists of the use of performance information for internal management decision-making (Binnendijk 2000). There are moves to ensure that monitoring and evaluation efforts are not merely limited to assessing performance, but also concerned with identifying and disseminating lessons learnt. The development of new forms of context-sensitive and process-based performance measurement systems that generate appropriate information is critical to performance improvement and organisational learning (Paton 2003).

PERFORMANCE MEASUREMENT IN PRACTICE: ITS EVOLUTION AND DEVELOPMENT

Historically, the role of performance measurement systems in the private, public and non-profit sectors has evolved over time. Theoretical notions of performance and efficiency were first introduced by Frederick Taylor at the beginning of the twentieth century. He proposed the application of scientific methods to discover the most efficient working techniques for manual forms of labour as well as to improve managerial control and supervision. Taylor's mechanistic management approach became known as 'Scientific Management'. Taylorism was seen to legitimise management as a control agent, and Taylor's techniques laid the groundwork for management control systems. Subsequent developments in the search for greater managerial efficiency were shaped by the work of highly successful managers, such as Henry Ford, in developing the early mass production systems; as well as the writings of theorists like Henri Fayol who identified the basic principles of rational administration. He defined the key managerial functions, including planning, organising, command, co-ordination and control (Hatch 1997; Courtney 2002).

The first decades of the twentieth century were marked by this positivist and mechanistic understanding of how human organisations function, and the belief in objective measurement (Lynch and Dicker 1998). New performance measurement systems were introduced. For example, the Dupont company developed a Performance Pyramid which linked financial ratios to rates of return on investment (Neely 2002). The growing awareness of the value of such systems and the recognition of the need for more sophisticated measures prompted organisations to

experiment with non-financial performance indicators to complement the range of financial measures already used. General Electric, in the 1950s, was one of the first organisations to implement a broad-based set of performance measures incorporating both financial and non-financial measures. The centrality of objectives to these processes was highlighted by Peter Drucker's work on Management-By-Objectives (Drucker 1973). Ongoing interest in performance measurement, and an increased understanding of the processes involved, led to a greater acceptance of the need to take a more holistic and balanced approach to such systems (Neely 2002). While, in the 1980s, much interest was centred on new measures of specific organisational dimensions designed to supplement financial measures, the focus of the 1990s was on integrated performance measurement frameworks such as the Balanced Scorecard and the Performance Prism. The involvement of key stakeholders in the processes is central to these frameworks; the way they identify and communicate their priorities and strategic needs is increasingly recognised as a crucial element in such systems (Neely 2003).

The integration of non-financial, qualitative, and process-based measures with existing financial and quantitative performance measurement systems reflects greater awareness of the complexity of the management process. Initially, organisations were depicted as mechanistic and rational entities that could be measured objectively; now they are increasingly perceived as evolving cultures and dynamic systems. There is currently greater understanding of the impact of human interaction and negotiation on organisational development and management processes. But despite this trend, the epistemological tensions between the quantitative and the qualitative, the hard and the soft, are inherent in many debates about effectiveness and efficiency.

The Evolution of Performance Measurement in the Private Sector

Measures of commercial success and performance such as profitability, sales turnover, and investment ratios have a long history in the business sector. However, as described earlier, most instrumental notions of efficiency and performance measurement did not appear until the beginning of the twentieth century. Performance measurement techniques and frameworks were further developed in the 1950s. This was exemplified by General Electric's identification of key corporate performance measures that placed an increased emphasis on quality and customer expectations (Eccles 1998).

In the 1980s, the quality movement entered the performance measurement scene in the private sector. Quality was seen as a strategic weapon in an increasingly competitive market place. Substantial resources were dedicated to developing measures such as defect rates, response time and delivery commitments in order to evaluate the performance of products, services and operations. By the 1990s, Total Quality Management had become a well-established tool to promote and assess improvements in quality. It gradually placed more emphasis on a range of qualitative factors and processes, and moved away from its original reliance on quantitative measurement (Czarnecki 1999). Metrics like quality, customer satisfaction, innovation, market share and human resource-related measures may be used to provide a more rounded picture of a company's condition and growth prospects than could be given by earnings ratios or other economic indicators. The 1990s also saw the development of competitive benchmarking techniques, which involve identifying competitors (as well as companies in other industries) who may represent best

practice in different activities, functions or processes – and comparing their performance with one's own (Eccles 1998).

The trend is therefore to combine different quantitative indicators and qualitative measures with processes that monitor the character and evolution of an organisation. These in turn are not static and can be adapted to changing priorities or organisational needs. For example, measures that were relevant during the early stages in the development of a new product may become inappropriate in later production stages. What is appropriate at the early stages of a change process may not be relevant as the organisation consolidates. Performance measurement indicators and methods are increasingly designed to suit the needs of individual organisations, and reflect the demands of the different markets in which they operate. Different businesses with different strategies require different types of information and analytical tools. Performance measurement is therefore no longer looked upon as a standard process with universal indicators. It is a strategic tool for management to increase productivity and profitability, and may help shareholders hold management accountable.

The Evolution of Performance Measurement in the Public Sector

The New Public Management strategies of the 1980s led to government efforts to improve the efficiency of public services. Concurrently, an unprecedented wave of interest in performance indicators arose, which has been associated with the New Managerialism and rational planning systems commonly found in the private sector of the 1960s (Carter 1994). The 1970s had seen some enthusiasm for Management-By-Objectives schemes in the public sector, and sporadic attempts were made at using target-linked performance indicators. In addition, social cost-benefit analysis was introduced as a way to quantify social effects, and attributing money values to non-monetary benefits. Thus, the quality of health and social services began to be measured with quantitative indicators. But it was in the 1990s that efficiency and effectiveness criteria were applied across the board to gauge performance and test whether public services were providing value for money. A new raft of performance indicators was introduced to allow government to control costs, increase accountability, and improve services (Rees 1992; Carter, Klein and Day 1992). The 'audit society' had arrived with a vengeance.

In the late 1980s and early 1990s, politicians and public sector managers set great store by performance measurement systems to deliver improved public sector services. The period saw a gradual shift from non-monetary measures towards quality-based effectiveness measures (Hyndman and Anderson 1998). The proliferation of customer care programmes shifted the assessment focus to the recipients of the services (Rees 1992). There was also an increased concern with the quality of provision as the concept of Total Quality Management spread among public services. But this shift to non-monetary measures has not lessened the public sector's preoccupation with indicator-driven management styles (Raine and Willson 1997). There has been an explosion in the quantity of performance indicators as well as a stronger emphasis on their role in quality improvement strategies. They cover such diverse concerns as waiting lists in hospitals, the performance of primary schools, the reliability of train services, or suicide rates in prisons. The use of performance indicators also reflects a growing tendency to audit activities in the public domain. However, there is concern that attempts to ensure the accountability of much of the public sector using such measures will prove complicated because of

the difficulties of attribution, the indirect lines of responsibility and the influence of politics and other external factors (Carter 1994). Public sector services are inherently interdependent and multi-dimensional and as a result a range of resource-intensive measures are needed to capture the full picture – all of which need to be co-ordinated through an overarching strategy (Sheldon 1998). For example, in the UK, performance frameworks are frequently developed in isolation from existing service frameworks, local quality assurance schemes or the needs of local councils – and, as such, they are therefore of little operational value. In light of this, it is understandable that there is growing criticism of the high number of indicators imposed by central government, and increasing pressure for governments to limit their reliance on inappropriate and irrelevant indicator-driven performance measurement systems.

The Evolution of Performance Measurement in the Non-Profit Sector

Performance in non-profit organisations first came into focus soon after World War Two when the extent of their performance problems was first exposed and scrutinised. The sector was increasingly criticised for its lack of professionalism and inability to meet performance targets (Courtney 2002). The 1960s and 1970s saw significant growth in the number of voluntary non-profit organisations. This was partly a result of public interest in the work of NGOs and other non-profits, and partly because governments saw these organisations as a convenient and inexpensive means of delivering public services and implementing development policies. In the 1980s, political pressure to reduce the role of the State, and a renewed emphasis on cost-effectiveness (notably in the UK and the US), led to a rapid expansion of the sector and the use of more formal contracts to define relations with non-profit organisations. They were seen as alternative service providers and drivers for improved performance in the newly competitive public sector. However, this contractual climate was marked by a New Managerialism with an emphasis on targets and performance indicators.

In the development sector, in which many non-profit NGOs currently operate, official donors have adopted the Logical Framework as their primary planning and performance assessment tool. This framework was originally created as a planning tool for military purposes, and was further developed by NASA to plan space programmes. It was subsequently adopted in the development sector by USAID in the early 1970s. European development agencies began to use the Logical Framework in the 1980s, and the German agency, GTZ, actively promoted ZOPP as an alternative version. ZOPP, or Objectives-Oriented Project Planning, encouraged a more participatory approach to objective setting and problem analysis (Helming and Gobel 1997). At the moment, Logical Framework Analysis is one of the key elements in defining the relationship between official donors and NGOs.

During the 1980s, 'the lost decade' as some would call it, questions were increasingly asked about the effectiveness of official aid, the quality of planning processes, and the capacity of MandE systems to judge the success of development interventions. Traditional top-down evaluation methods were criticised, and new bottom-up, participatory approaches developed. There was also a shift away from project work to more integrated long-term programmes. This involved the development of new techniques, and agencies such as DANIDA and DFID began to experiment with stakeholder analysis and more sophisticated participatory approaches, which they incorporated into their Logical Framework methodologies. Thus, there has been a change of focus from output to outcome and from product to process. This in turn has

led to an emphasis on qualitative indicators rather than the previous preoccupation with quantitative indicators.

Since the late 1990s, the focus has primarily been placed on issues of accountability, impact and effectiveness. The current evaluation debate has raised concerns about the methodological problems of measuring the effectiveness and impact of complex development programmes, as well as the performance of intrinsically complex aid interventions in areas such as advocacy and capacity building in unpredictable and rapidly changing environments (Cracknell 2002; Fowler 1997; Oakley, Pratt and Clayton 1998 and Oakley 2001). This has led to the development and application of alternative performance indicators and approaches (Hirschman 2002; Hailey and James 2003). It has also resulted in the development of new Results Based Management strategies, which attempt to integrate traditional performance measurement systems with broader programmatic goals (Binnendijk 2000).

It is apparent that a parallel evolution has taken place in the different sectors. While performance measurement was originally confined to financial and quantitative indicators, which were easy to define and measure, organisations in the three different sectors have begun to adopt more process-oriented approaches, which incorporate soft, qualitative and non-financial measures. While these more dynamic and flexible approaches may reflect reality to a greater extent, they also represent inherent methodological difficulties. There has similarly been a shift from merely perceiving performance measurement as a means of assessing success (or failure) to recognising its performance-enhancing role, as it highlights good practice and facilitates organisational learning. However, the wide variety of frameworks and methods currently applied to measure and manage performance situate themselves across the continuum between results-focused approaches and approaches emphasising learning for continuous improvement; the 'accountability-learning dichotomy' remains a cross-cutting tension.

SOME ALTERNATIVE APPROACHES

Performance measurement approaches are continuously criticised as insufficient and inappropriate. Therefore, various attempts have been made over the years to modify existing frameworks, use them in different ways, and create alternative frameworks. Often, the same framework or strategy is used for different purposes such as planning, management, and measurement, despite the fact that each process requires different framework features. Consequently, the terminology is not very precise; the same performance approach may be characterised by some as a performance management framework, and by others as a performance measurement system. Most approaches contain some form of measurable criteria or indicators, as well as one or more strategic dimensions or processes. These may be used for management and evaluatory purposes, respectively, depending on the role they are expected to perform.

In order to gain some insight into the diversity of tools and approaches currently being applied, the following section offers a brief analysis of the characteristics of some of the most commonly used performance measurement systems and management frameworks. These include: Total Quality Management, Benchmarking, the Balanced Scorecard, Performance Prism, Results Based Management, and the Logical Framework. While these are the most cited, we recognise that there are many other alternate audit frameworks such as Social and Environmental Audits,

publicly accredited standards such as ISO 9000, as well as government chartermarks and awards, such as Investors in People and consumer kitemarks.

Total Quality Management

Quality management is a widely used performance management strategy. Many versions have been developed over the years. The most renowned quality management strategy is Total Quality Management (TQM); a measurement and benchmarking strategy that creates an environment for organisational change by emphasising continuous improvement and the pursuit of best practice, and providing a customer and stakeholder focus (Czarnecki 1999). This quality management approach has been adopted by many organisations in the private sector, and more recently, in the public sector. TQM was introduced by Deming in Japan just after World War Two as a 14 point system for quality management. Its transfer to the US, and subsequently to the rest of the world, began in the 1980s inspired by Japan's superior performance in the industrial sector which was primarily due to its practice of quality improvement based on Deming's methods (Argyris and Schon 1996). After a surge of popularity in the 1980s, criticisms were raised during the early 1990s about TQM's limitations when managing complex organisations, and the ability of the strategy to shape individual performance levels was questioned. Furthermore, the assumption of direct causal linkages inherent in TQM only partially reflected the reality of performance management: it has proved inadequate in revealing specific performance gaps, and in identifying who is responsible for causes of process defects.

Total Quality Management was originally based around financial and statistical measures. Today, there is more focus on quality management processes as defined by the Malcolm Baldrige Quality Award criteria. These are based around seven categories: leadership, strategic planning, customer and market focus, information and analysis, human resource focus, process management, and business results. These categories are derived from eleven strategic core values, which include continuous improvement and learning, as well as a focus on results. For example, continuous improvement and learning is seen to be driven by the need for more responsive and efficient organisational processes, resulting in new and improved products and services. In recent years, a number of new quality performance approaches have been developed, including Six Sigma, a cost-saving, quality-focused strategy designed to reduce the number of defects and problems, which was developed by Motorola in the 1980s. However, many of them are criticised for being time consuming and resource expensive because they are inherently complex and overly procedural (Jarrar and Neely 2003). At the moment, TQM is a well-established and influential performance measurement technique. It takes a process-oriented approach which incorporates the views of key stakeholders, and emphasises innovation and improvement.

Benchmarking

Best-practice benchmarking is a process whereby organisations pursue enhanced performance by measuring their processes and practices against those in other organisations. For example, comparisons may be made with competitors, partners, and organisations with similar processes operating in different spheres, and even other parts of the same organisation (Neely 2002). Benchmarking is a strategy that can be used to measure performance and identify performance gaps, as well as

bring innovative ideas into organisational processes, thus improving organisational management (Czarnecki 1999). The shift to active organisational learning requires analysis of performance measurement by comparisons with better performing organisations, and investments of time and resources in order for improvements to take place (Letts, Ryan and Grossman 1999). The actual implementation of benchmarking varies with regards to the network of contacts created, the degree of mutual receptivity of ideas, and the willingness to test these ideas in a new context (Neely 2002).

Neely distinguishes between 'indicator benchmarking' which focuses on performance measurement, and 'ideas benchmarking' which is concerned with performance improvement. Benchmarking between similar organisations is likely to be competitive, and the process may be inhibited by considerations of commercial advantage. Consequently, similar benchmarking partners are more likely to pool indicators than share ideas. In the less competitive public sector, the number of benchmarking networks is increasing. Over the years, the notion of benchmarking has undergone an evolution from internal benchmarking to generic process benchmarking:

- *Internal Benchmarking*: comparing similar operations within one's own organisation
- *Competitive Benchmarking*: comparing with best practice of organisations in direct competition
- *Functional Benchmarking*: comparing practices and methods with similar processes in the same function outside one's industry
- *Generic Process Benchmarking*: comparing one's work processes with those organisations that have innovative, exemplar processes (Neely 2002, p.220)

Evidence suggests that there is a benchmarking maturity curve whereby organisations appear to move from simple comparisons of easily measured discrete activities with close partners, to comparing more complex processes with a range of partners. While indicator-driven benchmarking may lead to problems of comparability and consistency of data, the extensive use of comparative assessment of good (and bad) practice suggests that such an approach delivers immediate and operational benefits.

Balanced Scorecard

The Balanced Scorecard is a performance measurement and management framework which was developed by Kaplan and Norton in 1992 in order to measure the economic and operational performance of a firm. Since, the framework has been widely used by companies all over the world to track their financial results, while monitoring progress in building the capacities which they needed for growth (Kaplan and Norton 1998b;1998c).

The Balanced Scorecard aligns performance measures with organisational objectives through strategic performance management. The framework provides a balance between financial and non-financial measures; between internal and external

measures. It integrates four dimensions of performance, each containing goals and corresponding measures (Kaplan and Norton 1998a). These dimensions or perspectives are:

- *The Customer Perspective* – this refers to customer concerns about time, quality, performance, service, and cost.
- *The Internal Business Process Perspective* – this refers to those business processes that have the greatest impact on customer satisfaction.
- *The Innovation and Learning Perspective* – this refers to innovation, as well as improvements to existing products and processes.
- *The Financial Perspective* – this refers to the extent to which the company strategy contributes to profitability, growth and shareholder value.

Norton and Kaplan suggest that the Balanced Scorecard should be used to link long-term objectives with short-term actions – for example, assisting managers to clarify the organisational vision and build consensus among staff for corporate strategies, or helping organisations develop their strategic learning by facilitating dialogue and feedback. The Balanced Scorecard takes a systems perspective rather than a linear perspective, and has proved to be a successful tool for driving dynamic change (Ampuero, Goranson and Scott 1998). The framework links company strategy with the performance measures used to monitor and control strategy implementation. The Balanced Scorecard has primarily been used in the business sector, but it is increasingly applied by public and private non-profit organisations who aim to ensure ‘value for money’ and new efficiencies (Cutt and Murray 2000). While the four perspectives used in the Balanced Scorecard are relevant to all types of organisations, the choice of performance measures must be adapted to the specific goals and strategy of the organisation in question.

The Performance Prism

After analysing the weaknesses and characteristics of existing performance measurement systems, Neely (2002) developed his own performance measurement and management framework, the Performance Prism, which takes as its point of departure the reciprocal relationship with each stakeholder. The Performance Prism links value creation with performance measurement by investigating the organisational strategies, processes and capabilities needed to create value for stakeholders. The Prism is based on Neely’s belief that a good performance measurement system should provide a balanced picture of the business and its performance. He argues that any performance measurement framework should be multi-dimensional and comprehensive.

The Performance Prism is applicable across the organisation; across its functions and hierarchy; and entails methodologies for measuring specific activities or processes (Neely 2002). Compared with the Balanced Scorecard, it incorporates a range of stakeholder perspectives such as those of regulators and pressure groups. Neely presents five distinct but interlinked perspectives: stakeholder satisfaction, strategies, processes, capabilities and stakeholder contribution. The Performance Prism is designed as a dynamic model illustrating the interplay between these five perspectives. It can be applied at any organisational level, and provides a balanced

picture of the organisation, highlighting external (stakeholder) and internal (strategy, process, and capability) measures. It uses both financial and non-financial measures, and because it is multi-dimensional, it enables assessments of all the different influences on organisational performance. In a relief and development context, the Prism might complement the Logical Framework by taking an organisational perspective rather than an activity-based perspective.

Results Based Management

Results Based Management (RBM) is the approach currently adopted by most development agencies. It is a broad performance management strategy aimed at achieving changes leading to improved performance and better results. Performance indicators and the Logical Framework are key components of RBM systems. The RBM approach focuses on three key organisational levels: project level, country level and international level (or agency-wide). It combines strategic planning and performance measurement, and places these processes in a development context. Experience suggests that such processes are only fully effective if they can be applied at all levels of the audit trail, and data generation from all the different levels is carefully co-ordinated.

Commonly, RBM integrates such different elements as:

- the identification of objectives, indicators, and targets for each indicator
- the establishment of a monitoring system for regular data collection
- the comparison of targets with actual results
- the integration of evaluations to provide complementary performance information
- the use of performance information for internal management accountability, learning and decision-making processes, as well as for external performance reporting to stakeholders and partners (Binnendijk 2000).

The introduction of RBM systems reflects the increasing concern with the need to assess impact and ensure greater accountability. However, such systems may have a contradictory purpose, as organisations may try to use them as mechanisms for both accountability and performance improvement. While external stakeholders demanding accountability will tend to ask for results, internal stakeholders may be more concerned with how to improve the processes. Therefore, making the right choice about the role of the RBM system in specific cases is essential for good practice, as is recognising its limitations when attributing cause and effect, or using it as a means to apply sanctions and rewards (BOAG 2002).

The Logical Framework

The Logical Framework is virtually unknown outside the development community, and it is noteworthy that it has not been adopted to any great extent elsewhere. As has already been noted, it was introduced to the development sector by USAID in the early 1970s. The framework is built up as a linear model, based on the assumption of

objective causal relationships between inputs, activities, outputs, and objectives. It uses a matrix containing:

- a hierarchy of objectives for a specific project or an intervention
- indicators as to whether objectives have been achieved
- targets and sources of information for each objective level
- a set of assumptions concerning the preconditions for the desired project or intervention to succeed (Gasper 1999)

The Logical Framework is primarily used as a planning tool; a contractor's blueprint on which project funding is based, and by which performance can be monitored and evaluated. Its structured format is also prized as a communications tool – a mechanism for reducing the complex to the ordered. Moreover, it can be applied as a scheduling tool, and a checklist of activities and outputs, which in turn can act as an audit trail. While the intrinsic quality of the Logical Framework is its simplicity, experience has led practitioners to recognise that it needs to be complemented by different methodologies. Some development agencies, such as DANIDA, use the problem tree as a basis for stakeholder participation in goal setting. Others experiment with force field analysis or SWOT (strengths, weaknesses, opportunities and threats) analysis. CIDA has adapted the original framework and introduced the Results-Oriented Logical Framework (CIDA 2002b).

Despite its widespread use, there is a large body of criticism about the use of the Logical Framework as a development tool. Much of this criticism relates to its application rather than to the framework itself. It has been suggested that it is too rigid and mechanistic, and unable to reflect the complex reality of development projects implemented in unpredictable and culturally diverse environments. It is arguably an inflexible, top-down tool that reflects a bureaucratic, non-participatory view of the development process. It is also commonly seen as a blueprint that fails fully to reflect the complex processes involved, and its linear, overly structured design assumes causal links which may not reflect reality (Gasper 1999; 2000). Such criticism is only to be expected of a framework that was originally created as a planning tool for corporate and military purposes in the control and command planning culture of the 1960s. It should also be noted that much of the criticism of the Logical Framework could apply to any overly structured performance measurement framework, and this may explain the current interest in more process-based systems.

OPERATIONAL CHALLENGES

Questions about the appropriateness and credibility of Logical Framework and RBM strategies have led the development community to consider adapting some of the more innovative process-based elements found in other performance measurement systems. Yet, whatever approach is adopted, there are problems inherent in most kinds of performance measurement methodologies, because they provide a partial and incomplete picture of reality. Essentially, most criticisms relate to the application of performance measurement frameworks, rather than the frameworks *per se*. This section reviews some of the challenges faced by practitioners in trying to design and implement such frameworks. The value and credibility of frameworks is threatened by

a variety of methodological challenges, the costs involved, power imbalances, the influence of culture and context, and concerns about how to measure change in complex situations.

Methodological Challenges

While mechanistic or quantitative measures identify only parts of reality, relying exclusively on qualitative evidence equally represents methodological difficulties. There are concerns about the generalisability of such evidence, let alone worries that it is expensive to collect and interpret. We must also recognise that performance measurement and analysis is undertaken by individuals with subjective perceptions and personal agendas, who may have limited skills in this area. This will have a major impact on the way such measures are used and applied, and commonly leads to problems and even failure. Paton (2003) suggests that system failure is not only caused by problems with design and implementation, but also by measurement dysfunctionality – in other words, by intrinsic weaknesses inherent in performance measurement systems.

Such weaknesses are clearly apparent in attempts to use these systems as the vehicle for new learning and performance improvement. In the development sector, a common obstacle to organisational learning is constituted by the quantity of information, the quality of the analysis, the limited access to the data generated, and uncertainty as to the relevance of the material. Often, evaluation reports contain large amounts of complex and, at times, inappropriate information. It may therefore be difficult for different stakeholders to seek out and derive the information that applies to their particular needs. The question of attribution also represents a major obstacle to performance improvement. The interplay of internal and external factors as well as changing circumstances complicates the task of attributing responsibility and establishing causal relationships. If organisational learning and improvement is to take place, performance has to be measured in ways that allow changes and differences in performance to be attributed to particular contributions or conditions – a particularly difficult task when dealing with a multiplicity of actors and complex partnerships (Paton 2003; Engelhardt and Lynch 2003). There are clearly a wide range of methodological difficulties and process-related problems inherent in the way performance management systems are applied, which jeopardise their reliability and credibility. This then begs the question – what should the characteristics of such systems be, if they are to live up to stakeholder expectations, or what new systems might otherwise replace them?

Human Factors

Much of the success of such systems is determined by the human factor. The successful application of performance measurement systems depends on the intentions of those who commission it, as well as the resources, commitment and approach of those who design and operate it. The principal reasons why performance measurement initiatives fail are poor design of the measures themselves and difficulties related to the way the overall system is implemented. Managers tend to make fundamental mistakes when choosing what to measure, and selecting the performance indicators. Some performance measures are inappropriate and impractical in an organisational setting. Furthermore, the way measures are implemented often alienates staff and volunteers who feel threatened by the indicators used or the processes initiated. They associate them with sanctions or see

them as bureaucratic control mechanisms, and so try to manipulate data and present results in a favourable light – all of which further undermines the credibility and value of such systems. Staff are also critical of the time and resources needed to gather and analyse data, and the quantity of paper generated to support the audit trail. The overall process is time consuming, frustrating, and deflects staff from their primary tasks. Consequently, it is understandable that staff and volunteers often do their best to derail and subvert such measurement processes.

Lack of Investment

The human alienation referred to above also reflects the general lack of investment in measurement and change processes. Such processes are expensive. They depend on appropriate institutional capacity as well as considerable investment of time and effort. Often, insufficient funds are made available, too limited time allocated, or too weak infrastructures are created to collect and analyse performance data satisfactorily. Furthermore, the organisation's priorities may shift during the process, and the infrastructure needed to support the performance measurement system may become neglected. Finally, there is a tendency to introduce new 'vogue' indicators, and such measurement manoeuvres place unnecessary pressure on existing infrastructure, further disengage staff, and deflect focus away from well-established, long-term indicators of performance (Neely 2003, Neely and Bourne 2003). Many of the problems and dysfunctions outlined above have led to a review of the way performance indicators are used. For example in the UK, the Chief Inspector of Schools and the Home Secretary voiced their fears in March 2003 that the use of indicators as a way of enforcing government policy was increasingly counter-productive. They both concluded that it had alienated those involved and led to inappropriate practices.

Transaction Costs

The increased use of performance management systems comes with a cost. The direct and indirect costs attached to introducing and applying such complex systems are high. They are expensive in terms of time invested, associated staff and consultancy costs, data collection and presentation – and in addition there are a range of indirect opportunity costs, and the cost to the organisation provoked by personal frustration, resistance and low morale. In most NGOs, such costs must be financed out of overheads, and commonly result in creating additional budget constraints.

The motivation among NGO staff to engage in performance measurement may depend on whether they can justify these costs in terms of greater transparency, personal learning or improved performance (Letts, Ryan and Grossman 1999). Moreover, the development of process-oriented participatory approaches, which depend on full stakeholder involvement on equal terms, demands additional investment in capacity building, information sharing and better communication. It is therefore crucial that donors appreciate the full costs involved in the application of effective and appropriate performance measurement systems, and allocate sufficient funds to cover the direct and indirect costs associated with such systems.

Power and Control

There is an undercurrent of concern that performance measurement systems are manifestations of donor power or managerial control. The degree of control and power exerted and accepted by different stakeholders may differ, as not all stakeholders are equally influential. Therefore, negotiations around how to prioritise indicators and performance measures are likely to be unbalanced. NGOs frequently attempt to overcome this by encouraging a culture of consensus and including stakeholders in negotiations, but this in turn creates problems, as such efforts tend to cover up informal power imbalances (Paton 2003). Influential stakeholders, donors not least, are often able to exert power over organisations, and hence, tighten the control. These stakeholders need to acknowledge and address the impact of such power differentials on the measurement process as openly as possible. On the other hand, experience suggests that they are hesitant to impose sanctions, if terms and conditions are not met, for fear of a backlash of accusations as to their abuse of power. As a result, they too are disempowered and open to manipulation. Furthermore, competition between donors to allocate aid to particular groups in priority areas weakens their influence. Other factors, such as political imperatives or development fashions, may also affect the ability of donor representatives to take an objective stance and stop support to failing programmes. In some circumstances, donors may actively preserve programmes that are not objectively successful; for example, evidence from Uganda suggests that some failing projects continue to be given more favourable support, because Uganda is a country favoured by donors (Adam and Gunning 2002).

Culture and Context

Globalisation has led to widespread interaction between actors from different cultures and contexts. At this point in time, a plethora of multilateral agencies and international NGOs operate around the world. Cultural differences impact on the design and operation of performance measurement systems. In addition, the choice of performance measures will be determined by a range of cultural and contextual factors. Not only is there a need for greater understanding of the impact of different cultures on the choice and use of performance measures, but we need to recognise how differing cultural values impact on the application of different assessment processes (Hirschmann 2002). For example, culture may influence the way such processes are seen and the weight given to different performance measures. There is evidence to suggest that organisations operating in high power distance cultures or ones with high levels of uncertainty avoidance will be more likely to adhere to the idea of performance measurement and give more credence to what these measures represent (Hofstede 1991). Cultural and contextual differences may also limit our ability to compare the performance of similar organisations working in different cultures and settings, and hence undermine any comparative analysis.

Complexity and Change

Complexity, uncertainty and the impact of ongoing change have a considerable impact on organisational performance. This is partly reflected in the diversity of relevant stakeholder relationships, and the level of interaction between individuals and organisations. Attempts have been made to face this challenge by introducing more complex frameworks and increasing the range of performance indicators used. But such action comes at the expense of simplicity, functionality, and accessibility.

This tension is becoming even more significant with the growing investment in such intangible, ill-defined processes as capacity building, advocacy, and community empowerment. The challenge for performance management approaches is how best to assess these complex activities against concrete indicators, and install measurement processes that satisfy stakeholder demands for accountability and still make performance improvement possible. Uncertainty and change are interlinked. Performance measurement processes must be regularly re-evaluated and new indicators introduced, in order to assess the degree of change taking place and make the appropriate adjustments. Davies (2002) addresses this issue by analysing different approaches to measuring change. He describes change as the result of reciprocal influences in a development process, and suggests that many parallel change processes may take place concurrently. He argues that the Logical Framework, with its rigid structures, is unable to capture the true nature of change; more flexible approaches, such as process mapping, are needed to overcome the problems associated with monitoring projects or programmes subject to rapid change and conflicting interests (as is often the case in humanitarian operations). Development is inherently about change, and in particular, changes in the way individuals and communities relate to each other. Consequently, performance measurement should focus on changes and shifts in relationships rather than the causal linkages between objectives and outcomes (Davies 2002; CDRA 2002).

We have seen that there are a number of factors that impact on the efficacy of performance management systems. These factors include the impact of culture, the methodologies used, the costs involved, or the appropriateness of such measures in times of change, and affect the way performance measurement systems are perceived and applied. It should be recognised that most of these issues are interlinked: change contributes to complexity; control issues may lead to demands for accountability; cultural differences may affect power relationships, or the motivation for organisational learning and innovation. It would appear that stakeholder involvement is critical to deal with all issues. Meanwhile, increased stakeholder participation and the introduction of more culturally appropriate and process-based indicators will impact on the level of investment needed to cover the costs of the physical, financial and human resources involved. Faced with these challenges, a number of trends can be identified that are likely to shape the future development of performance measurement systems. We explore these trends below.

TRENDS IN THE DEVELOPMENT SECTOR

The focus on performance measurement in the development sector as well as the shortcomings of commonly used methods in this area have led to increased experimentation with alternative frameworks and strategies. Below, we present a brief account of the innovations and trends that are shaping the development of new performance measurement systems.

Participatory Performance Measurement

Participatory performance measurement involves greater stakeholder participation in the development of indicators as well as the actual measurement process and associated analysis. Agencies are experimenting with a variety of different strategies and frameworks which actively involve key stakeholders at all levels. For example, INTRAC's Central Asia programme is currently experimenting with a version of this approach, linking measurement to capacity building and goal-setting. Stakeholders in

the local community have actively participated in the formulation of objectives, the development of indicators, and are currently undertaking data collection based on these indicators. Initial conclusions about the process suggest that their involvement has meant that they are more motivated, more engaged in the processes, and more open to learning. However, the process is very time-consuming, and it has been hard to maintain a consistent level of participation. Varying degrees of stakeholder involvement at times increased the risk of imposing INTRAC's preconceived ideas about performance measurement. Furthermore, the participants' poor analytical capacities raised questions about whether the participants should be involved in the analysis of the data generated. This in turn has resulted in calls for greater investment in analytical skills development and wider capacity building and training efforts, in order to improve the participatory process.

There are other examples, where donor agencies have tried to involve local communities in assessing the effectiveness of broad-based development processes. USAID has set up an advocacy index, as a way of measuring performance of advocacy activities undertaken by local NGOs. The approach combines stakeholder involvement in the development of indicators with evaluations conducted by an independent panel of local consultants (based on these indicators). The purpose of the index is to quantify qualitative performance measurement in order to make comparisons possible between different NGOs. The approach is currently being tested in Zimbabwe and has proved successful in involving and motivating local communities – but it is still a struggle to attribute the responsibility for the direct consequences of any advocacy initiative, because of the difficulties linked to identifying the causal relationship between overall policy change and specific advocacy programmes (Hirschman 2002).

Growing Consensus on the Need for Impact Assessment

As we have seen, donor expectations and greater public scrutiny have led to greater demands for increased accountability as to the way public funds are spent, and for evidence of resulting impact. Aid agencies are under continuous pressure to demonstrate the benefit of development assistance by assessing its long-term impact rather than merely identifying the activities and outputs that may have occurred. The ongoing debate about the primary purpose of assessing impact has uncovered a number of different interpretations of the role of impact assessment systems, including:

- *The Legacy Role:* identifying what changes have occurred, analysing and documenting them by recording these changes and their impact.
- *The Communication Role:* communicating achievements internally and celebrating long-term successes in a way that motivates staff and stakeholders and encourages others to adopt and invest in similar processes (disseminating 'evidence of impact' is also a good PR tool that may help raise external funds or sell one's services).
- *The Governance Role:* ensuring a degree of accountability by monitoring investment into specific activities and outputs, and tracking the outcomes and impact in a systematic and transparent manner.

- *The Learning Role*: generating information and perspectives on the changes which have occurred, and analysing and disseminating them in such a way that all stakeholders can learn from the relationships and processes involved; from successes as well as failures.
- *The Policy Role*: generating data and analysis that can be used to reform policies and develop new strategies, improve government and donor practices, or strengthen NGO advocacy campaigns and efforts to influence policy-making.

Evidence from the field confirms that those involved in designing impact assessment systems need to distinguish these different roles to ensure that data is collected, analysed and disseminated in ways that support the particular role an assessment is expected to play. The choice of role will be closely linked to the primary purpose of the assessment in question.

There appears to be greater consensus about the key characteristics of effective impact assessment systems. These include stakeholder participation, self-evaluation, triangulation, and the use of a mix of methods and tools as well as solid analysis and follow up through dissemination of findings. The implications of including these elements in an impact assessment process are that successful assessment is costly and dependable on significant financial investments (Hailey and James 2003). Donors need to cover the direct and indirect management costs involved. If this is not done, the assessment will be poorly conducted, generate little information of operational value, and suffer from limited credibility. Impact assessment, alongside other forms of performance measurement, is likely to shape our thinking about the way aid should be provided in the future. It increasingly dictates the procedures and processes by which official aid is dispensed, and it is crucial that adequate investments are made to ensure the quality and reliability of such assessments.

Monitoring and Evaluation in the Humanitarian Context

With the increased intensity of humanitarian action in the late 1980s and throughout the 1990s came the realisation that monitoring and evaluation (M&E) approaches which were created for development purposes needed some adjustment to be applied in relief and emergency situations. As most emergency situations are characterised by extreme urgency and a multitude of actors, many of the operational challenges mentioned previously are accentuated. For instance, circumstances tend to change rapidly, and this creates a need for continuous monitoring which is often difficult to satisfy, as most field practitioners are tied up in operations and do not perceive monitoring as a priority. Many evaluations are conducted after the interventions have taken place, such as After Action Review (AAR). AAR focuses on a few key questions in order to obtain a quick picture of the process and outcome of the intervention. However, the tendency to link humanitarian action more closely with rehabilitation, reconstruction, and longer-term development raises demands for more comprehensive systems that can contribute to organisational learning as well as operational improvements. The organisational dimension tends to be left out of humanitarian evaluations which in most cases focus on the interventions. In addition, humanitarian evaluations have tended to be top-down, and to concentrate on the effectiveness of operations rather than the impact for beneficiaries.

Emergencies are becoming increasingly complex with elements of conflict, famine, forced migration, and a large number of local and international stakeholders with different (often opposing) interests, together with a range of interventions by various

agencies. Although some procedures and systems need to be in place in order for agencies to act quickly, humanitarian responses must not be too standardised, but carefully designed for the specific case. Otherwise, there is a danger that the underlying local processes of social change are undermined by such standard interventions. Monitoring and evaluation during (and not just after) the interventions will help ensure that adjustments are made as needed. As information sharing and co-ordination of all emergency operations taking place in a particular context is key to successful outcomes, initiatives are increasingly being taken to conduct evaluations that include the interventions of several agencies.

The complexity and instability in many emergency situations complicates the task of conducting a balanced analysis taking into account contributions from all stakeholders, and regional variations. Furthermore, access to relevant information and stakeholder participation in the M&E process is impeded by the weak capacity and frequent displacement and traumatisation of target populations caused by the insecure conditions they are facing. This implies that appropriate M&E systems for humanitarian assistance should be:

- *Solid*: allowing for structured, systematic and consistent processes to ensure data reliability and to institutionalise the necessity of M&E in humanitarian action.
- *Flexible*: adjusting to rapid changes and varying contextual factors, and allowing for different methods to be applied as appropriate when interacting with the diverse range of stakeholders.
- *Pragmatic*: relying on qualitative information and narratives rather than quantitative data collection which is likely to be incomplete.
- *Participatory*: contrary to the notion of 'victims' with limited agency and control which has been predominant until recently, target populations should be recognised as social actors with a capacity to influence the situation and initiate changes. To the extent possible, beneficiaries and other stakeholders in the local context should be included, although power imbalances may cause fear and significantly influence the statements given by some participants.

Measuring Change Produced by Capacity Building Interventions

It is generally recognised that the institutions and organisations which are at the forefront of the development process are so over-stretched and under-resourced that operational effectiveness suffers, and organisational viability and sustainability is threatened. As a result, it is crucial to encourage efforts to build the capacity of such organisations. There is an increasing understanding of the dynamics and efficacy of different capacity building interventions as well as an appreciation of the need to assess their long-term impact. However, it is clearly difficult to attribute the benefits of an investment in managerial functions and processes – such as establishing the causal relationship between the installation of a new accounting system in a clinic and the resulting improvements in the lives of the most vulnerable, or uncovering the link between a new or modified strategy at head quarters and actual changes in field operations. All too often, investments in organisational capacity take time to impact on the quality of work or the services offered, and efforts to link investment in new capacity with higher quality of life are fraught with methodological difficulties.

The challenge for those involved in assessing the value of capacity building interventions is great: first, they need to generate sufficient base-line data on which to gauge the quality and depth of any changes which may have occurred – in particular, they must obtain sufficient oral accounts and graphic images to portray these changes and track their impact on the lives of local people. Secondly, they should be able to identify direct causal links between a particular capacity building intervention and wider developmental benefits. In this respect, any capacity building intervention is like a drop of rain or a pebble that lands in the water – the ripples may flow outwards from changes brought about at an individual level to the organisational, inter-organisational and societal levels (James 2002). Thus, it is plausible that a training course may lead to improved knowledge as well as new skills and attitudes at an individual level. If course participants have been able to implement this learning in their organisations, it may mean that change will be experienced at a wider organisational level. Furthermore, better technical skills of staff and enhanced functioning of the internal organisation should lead to improved quality of the services provided to communities. This ought in turn to result in changes in the situation of the ultimate beneficiaries.

The size and direction of the ripple is influenced by (and in turn influences) the context in which it moves. Thus, interventions at the organisational level may lead to better conditions for staff development or inter-organisational relationship building; changes in the enabling institutional environment may lead to improved financial conditions for the individual organisation and enhanced communication between organisations; and the encouraged participation of beneficiaries (as primary stakeholders) in organisational decision-making processes may lead to the implementation of more relevant development activities by the organisation in question.

It is clear from this ripple model that any one particular capacity building intervention is merely one intervention among many that bring about change. Institutional changes, initiatives at community level, or development interventions by other agencies in the same area are bound to affect the changes induced by a particular capacity building intervention. But by using the concept of plausible association, it is possible to judge whether change at one level does indeed ripple out to bring about changes at a wider level. This, in turn, depends on the base-line data available for assessments of the quality and significance of the changes which have taken place. Consequently, there is a new awareness of how investing in collecting and analysing such data can add to our understanding of the dynamics of change. By applying concepts such as plausible attribution, these processes may help demonstrate the link between a particular capacity building intervention and its wider developmental benefits.

Measuring Core Values

In the non-profit and NGO sector, it is increasingly recognised that NGOs should consider compatibility with their core values, organisational identity and mission, when negotiating performance measures. According to Gasper (1999) NGOs are not merely delivery channels implementing externally imposed strategies; staff and members are motivated by their own vision, values and beliefs. Hailey (2000) also stresses the need for NGOs to develop their distinctive core values, in order to preserve their roles as significant development actors. He draws attention to the risk that the increased competition for limited aid funds will allow donors to impose conditions and influence the core values of NGOs. Hailey therefore proposes the

introduction of key 'indicators of identity' in order to assess the values and organisational capacities that distinguish NGOs from other agencies. He suggests that such value-based indicators should be fully integrated in organisational evaluations alongside other performance measures. Awareness of and commitment to core values may enable NGOs to negotiate more strategically with stakeholders, particularly when undergoing or undertaking performance measurement reviews.

There is a growing body of research that reflects our greater understanding of the role and influence of core values in creating a viable and vibrant organisational culture in both the private and non-profit sectors. Research into the management of NGOs in South Asia highlights the influence of such core values in facilitating change and motivating staff (Smillie and Hailey 2001). In the private sector, firms such as Hewlett Packard have developed corporate strategies based on core values including trust and respect, integrity, teamwork, flexibility and innovation. These companies actively share and sustain such values by investing in direct strategic human resources interventions (Letts, Ryan and Grossman 1999).

Culture and Context-Specific Performance Measures

There is currently greater awareness of the challenges of transferring models and methodologies to different cultures and contexts. Too many development projects, which looked good on paper, have failed to achieve projected results due to the impact of local circumstances. One consequence of this is increased investment in research and capacity development to facilitate our understanding of the impact of culture, or assess the extent to which capacity building methodologies are culture-bound. This is exemplified by the recent funding from the Dutch Ministry of Foreign Affairs for INTRAC's new capacity building programme, Praxis, which investigates transferability as one of its principal themes. The evidence suggests that culturally sensitive processes and relevant indicators need be developed to meet the demands of different settings. The transferability of performance measurement systems depends on how far they can be adapted to local circumstances while remaining internationally accepted and comparable.

One particular trend of note is the way in which new performance measurement frameworks are being specifically designed for the needs of particular organisations, and reflect the culture in which these organisations operate (Kelly 2002). While such tailor-made frameworks and context-specific measures can reflect the exigencies of the local culture and the needs of local communities, they are expensive to design and of limited value in making international comparisons. The key lies in the appropriate application of performance measurement methods. This relies as much on the intercultural competencies of the individuals conducting the assessment as on the relevance of the adjustments they choose to make of an existing measurement system in order for it to work in the local culture and context. It is necessary to compromise between the need for comparability for future lesson learning at a more general level, and the wish to track the actual changes resulting from an intervention, which demands an overview and an in-depth understanding of the specific situation.

Incorporating Complexity in Performance Measurement

New performance measurement systems are continuously proposed which attempt to cater for greater complexity and its implications. These rely on the use of a range of sophisticated indicators and new integrated, multi-dimensional and multi-level

frameworks. Other innovations concentrate on methods rather than frameworks: Cracknell (2002) suggests that the future might lie in a combination of participatory approaches and methodological pluralism. Confusingly, the development of these new integrated systems and complex multi-purpose frameworks comes at a time when there are increased calls for simpler and more easily applicable methods to assess the success of development interventions. Many development practitioners working in the field are calling for the use of fewer, essential, and better designed indicators, and the Development Assistance Committee (DAC) encourages the development of relatively simple and user-friendly performance measurement systems (Binnendijk 2000). These contradictory trends well reflect the tensions between evaluators and evaluated. Most designers of performance management systems have different agendas from practitioners in the field. Donor demands ought to be balanced with the needs of beneficiaries or the priorities of programme managers, when evaluations are conducted or impact assessments undertaken.

One consequence of this methodological dilemma is that a process of triangulation is increasingly used to analyse complex programmes. Triangulation is achieved by using a mixture of methods, tools and perspectives. This helps balance subjective and carefully delimited data collection with a certain degree of objectivity obtained by using multiple data sources and working in co-ordinated assessment teams (Adams 2001). Triangulation represents a pragmatic way of measuring complex realities using both quantitative and qualitative methodologies, while aiming to produce valid measurement results and assessments credible to all stakeholders. The evaluation team should include all stakeholder perspectives in a more balanced way, to ensure stakeholder accountability in its broader sense rather than prioritising the donor side. This approach also tends to facilitate learning at a wider level: in this way, the findings and recommendations arising from the evaluation process will concern many different actors involved in similar processes.

Understanding Change through Performance Measurement

Change, like complexity, has led to the development of new ways of working with performance systems and strategies. The Performance Assessment Resource Centre (PARC) and DFID are currently assessing the role of change forecasting as a way of collectively articulating potential changes and making them explicit to all stakeholders (Engelhardt and Lynch 2003). One way to assess the progress of change is to use 'milestones' of change. The milestones method represents a relatively simple way to monitor whether short-term tactical objectives have been achieved (Cracknell 2002). The use of milestones makes it easier to cope with elements of uncertainty, and helps gauge the extent of potential problems and the risks involved. New experimental approaches to assessing the extent of change and the dynamics of changing relations have attracted considerable interest. For example, narrative is currently being used in innovative ways by some agencies to illustrate the impact of change. The South African NGO, CDRA, suggests that story-telling is a useful means of assessing the impact of social development interventions; such outcomes are best uncovered by examining shifts in relationships which may reflect increased empowerment. Beneficiaries are asked to tell stories from which the organisations that undertook the interventions subsequently derive information about changes that have taken place (CDRA 2001).

While it is clear that it is important to measure change, performance measurement systems themselves should also evolve over time (Paton 2003). In the short-term, such systems may be adjusted to changing circumstances and managerial priorities by changing performance indicators, ensuring continuous monitoring or adopting the

milestones approach. In the longer-term, performance measurement systems and frameworks should be modified to reflect changing organisational imperatives and societal concerns.

CONCLUSIONS

This paper has offered an overview of some of the issues and innovations in the area of performance measurement and management. It is clear that this is a rapidly changing field, which is beset by definitional problems, methodological debates, contradictory criticisms, and uncertainty as to the primary purpose of performance management systems or strategies. We conclude that the key is not the choice of framework, but how it is applied. It is essential to make the process as important as the product; to create living frameworks that reflect the dynamic and multi-dimensional character of what they are trying to measure. Moreover, there are a number of operational challenges and crucial issues that need to be addressed to ensure the successful application of such performance measurement systems, notably around how best to promote learning and performance improvement while preserving a degree of accountability and transparency. In addition, issues related to how best to factor in power and control, culture and context, as well as complexity and change, must be considered. There is ongoing concern about the cost of implementing such systems; whether or not they are cost-effective and deliver what they purport to.

Many of these issues can be addressed through greater stakeholder participation in the development of objectives and performance indicators; in their analysis as well as the dissemination of findings. While such participation is time-consuming and expensive, stakeholder involvement is crucial, if the ultimate purpose of performance measurement, namely performance improvement, is to be achieved. It is therefore apparent that dedicated resources in terms of time and money are needed to make such stakeholder involvement possible. All the evidence suggests that successful evaluation and measurement strategies depend on significant financial investments. The transaction costs of performance management frameworks are high, and if they are to be applied successfully, donors need cover the direct and indirect management costs and associated overheads. If this is not done, performance measurement processes will be poorly managed and generate little information of operational value. Moreover, the findings will suffer from limited credibility. If donors are not prepared to cover the full cost of effective and appropriate measurement processes, they must question whether they should continue to impose them. The minimalist approach to project funding with its emphasis on trust and low transaction costs is a viable and cost-effective alternative that ought to be considered. One can only conclude that unless performance measurement systems are reliable, credible, and trusted, they have little role in measuring success.

BIBLIOGRAPHY

Adam, C. S. and Gunning, J. W. (2002) 'Redesigning the Aid Contract: Donors' Use of Performance Indicators in Uganda', *World Development* 30 (12): 2045–56.

Adams, C. and Neely, A. (2000) 'The Performance Prism to Boost M and A Success', *Measuring Business Excellence* 4 (2):19–23.

Adams, J. (2001) *NGOs and Impact Assessment*, NGO Policy Briefing Paper No.3, Oxford: INTRAC.

Ampuero, M., Goranson, J. and Scott, J. (1998) *Solving the Measurement Puzzle – How EVA and the Balanced Scorecard Fits Together*, The Cap Gemini Ernst & Young Center for Business Innovation. Issue 2 Measuring Business Performance: 45–52.

Argyris, C. and Schon, D. A. (1996) *Organizational Learning II – Theory, Method and Practice*, Upper Saddle River, NJ: Addison-Wesley Publishing Company.

AUSAID (2002) *The Logical Framework Approach*, in AusGUIDE, AUSAID website (www.usaid.gov.au), visited March 2002.

Bakewell, O., Adams, J. and Pratt B. (2003) *Sharpening the Development Process – A Practical Guide to Monitoring and Evaluation*, Oxford: INTRAC.

Bean, J. and Lascelles H. (1998) *Quality in the Public Sector – Essential skills for the Public Sector*, London: HB Publications.

Bell, S. (2000) 'Logical Frameworks, Aristotle and Soft Systems: A Note on the Origins, Values and Uses of Logical Framework, in Reply to Gasper', *Public Administration and Development* 20 (1): 29–31.

Binnendijk, A. (2000) *Results Based Management in the Development Co-operation Agencies: A Review of Experience*, executive summary of draft report, DAC Working Party on Aid Evaluation.

BOAG Evaluation Group (2002) Performance Assessment Seminars 4, July.

Boyne, G. A. (2002) 'Public and Private Management: What's the Difference?' *Journal of Management Studies* 39 (1): 97–122.

Brignall, S. (2002) *The Unbalanced Scorecard a Social and Environmental Critique*, Best Award Paper for the Performance Measurement Award 2002.

Canadian International Development Agency (2002a) *Assessing Sustainability*, in What We Are Learning series, No.2, Performance Review Branch, August 2002.

Canadian International Development Agency (2002b) *The Logical Framework: Making it Results-Oriented*, CIDA website (www.acdi-cida.gc.ca), visited December 2002.

Carter, N. (1994) 'Performance Indicators: 'Backseat Driving' or 'Hands Off Control'', in: McKeivitt, D. and Lawton A., *Public Sector Management – Theory, Critique and Practice*, London: Sage Publications.

Carter, N., Klein, R. and Day, P. (1992) *How Organisations Measure Success – The Use of Performance Indicators in Government*, London: Routledge.

Community Development Resource Association (2001) 'Measuring Development – Holding Infinity', Annual Report 2000/2001.

Courtney, R. (2002) *Strategic Management for Voluntary Nonprofit Organisations*, London: Routledge.

Cracknell, B. E. (2002) *Evaluating Development Aid – Issues, Problems and Solutions*, London: Sage Publications.

Cutt, J. and Murray V. (2000) *Accountability and Effectiveness Evaluation in Non-Profit Organisations*, London: Routledge.

Czarnecki, M. T. (1999) *Managing by Measuring – How to Improve Your Organization's Performance Through Effective Benchmarking*, Houston: The Benchmarking Network.

Davies, R. (2002) 'Improved Representations of Change Processes: Improved Theories of Change', paper presented at the 5th Biennial Conference of the European Evaluation Society, Seville 2002.

Drucker, P. F. (1973) *Management: Tasks, Responsibilities, Practices*, New York: Harper and Row.

Drucker P. F. (1990) *Managing the Non-Profit Organisation*, Oxford: Butterworth Heinemann.

Harvard Business Review (1998) *Harvard Business Review on Measuring Corporate Performance*, Harvard Business Review Paperback Series, Boston: Harvard Business School Press.

Earle, L. (2003) 'Lost in the Matrix: The Log Frame and the Local Picture', paper presented at the INTRAC 5th Evaluation Conference, Kontakt der Kontinenten, the Netherlands, March 2003.

Eccles, R. G. (1998) 'The Performance Measurement Manifesto', in *Harvard Business Review*.

Engelhardt, A. and Lynch R. (2003) 'Moving Upstream – New Challenges for M and E', SPARCS, Newsletter from the PARC, Performance Assessment Resource Centre, Spring 2003.

Frigo, M. L. (2003) *A Balanced Scorecard Framework for Internal Auditing Departments*, The Institute of Internal Auditors Research Foundation, Florida 2003, editorial summary by The Institute of Internal Auditors, www.iaa.org.

Fowler, A. (1997) *Striking a Balance – A Guide to Enhance the Effectiveness of Non-Governmental Organisations in International Development*, London: Earthscan.

Gasper, D. (1999) 'Problems in the Logical Framework Approach and the Challenges for Project Cycle Management', *The Courier* 173.

Gasper, D. (2000) 'Evaluating the Logical Framework Approach towards Learning-Oriented Development Evaluation', *Public Administration and Development* 20 (1):17–28.

Hailey, J. and James, R. (2003) 'NGO Capacity Building: The Challenge of Impact Assessment', paper presented at the New Directions in Impact Assessment for Development conference, University of Manchester, November.

Hailey, J. (2000) 'Indicators of Identity – NGOs and the Strategic Imperative of Assessing Core Values', *Development in Practice* 10(3): 402–7.

Hatch, M. J. (1997) *Organization Theory – Modern Symbolic and Postmodern Perspectives*, Oxford: Oxford University Press.

Helming, S. and Goebel, M. (1997) 'ZOPP – Objectives-Oriented Project Planning – A planning guide for new and on-going projects and programmes', GTZ (Deutsche Gesellschaft für Technische Zusammenarbeit), Unit 04 (Strategic Corporate Development), Eschborn.

Hirschman, D. (2002) 'Implementing an Indicator: Operationalising USAID's Advocacy Index in Zimbabwe', *Development in Practice* 12 (1): 20–32.

HM Treasury (2001) 'Choosing the Right Fabric: A Framework for Performance Information', London: Cabinet Office, National Office, Audit Commission, Office for National Statistics.

Hofstede, G. (1991) *Cultures and Organisations, Software of the Mind*, London: McGraw Hill.

Hubbard, M. (2001) 'Shooting the Messenger: Log Frame Abuse and the Need for a Better Planning Environment – a Comment', in *Public Administration and Development* 21 (1): 25–6.

Hyndman, N. S. and Anderson, R. (1998) 'Performance Information Accountability and Executive Agencies', *Public Money and Management* 18 (3): 23–30.

James, R. (2002) *People and Change – Exploring Capacity Building in NGOs*, Oxford: INTRAC.

Jarrar, Y. and Neely, A. (2003) 'Six Sigma – Friend or Foe?', Centre for Business Performance, Cranfield School of Management.

Kanter, R. M. and Summers, D. V. (1994) 'Doing Well While Doing Good: Dilemmas of Performance Measurement in Nonprofit Organizations and the Need for a Multiple Constituency Approach', in D. McKevitt and A. Lawton *Public Sector Management – Theory, Critique and Practice*, London: Sage Publications.

Kaplan, R. S. and Norton D. P. (1998a) 'The Balanced Scorecard – Measures that Drive Performance', in *Harvard Business Review* (originally 1992).

Kaplan, R. S. and Norton, D. P. (1998b) 'Putting the Balanced Scorecard to Work', in *Harvard Business Review* (originally 1993).

Kaplan, R. S. and Norton, D. P. (1998c) 'Using the Balanced Scorecard as a Strategic Management System', in *Harvard Business Review* (originally 1996).

Kelly, L. (2002) 'International Advocacy: Measuring Performance and Effectiveness', paper presented at the Australasian Evaluation Society International Conference, Australia October/November 2002.

Knapp, M. (1996) 'Are Voluntary Agencies really more effective?', in D. Billis and M. Harris (eds) *Voluntary Agencies – Challenges of Organisation and Management*, London: Macmillan.

Lawrie, A. (1992) *Quality of Service – Measuring Performance for Voluntary Organisations*, London: Directory of Social Change/National Council for Voluntary Organisations.

Leat, D. (1999) 'Voluntary Organizations and Accountability: Theory and Practice', in M. Anheier and W. Seibel de Gruyer (eds) *The Third Sector – Comparative Studies of Nonprofit Organisations*, Berlin: Walter de Gruyter.

Leat, D. (1996) 'Are Voluntary Agencies Accountable?', in D. Billis and M. Harris *Voluntary Agencies – Challenges of Organisation and Management*, London: Macmillan.

Leeuw, F. L. (2002) 'Assumptional Analysis, Log Frame Analysis and Other Methods of Reconstructing and Evaluating Program Logic', Working Paper, Utrecht University.

Letts, C. W., Ryan, W. P. and Grossman, A. (1999) *High Performance Nonprofit Organizations – Managing Upstream for Greater Impact*, New York: John Wiley and Sons.

Lindenberg, M. (2001) 'Are We at the Cutting Edge or the Blunt Edge? Improving NGO Organisational Performance with Private and Public Sector Strategic Management Frameworks', *Nonprofit Management and Leadership*, 11: 247–70.

Lynch, T. D. and Dicker, T. J. (1998) *Handbook of Organisation Theory and Management – The Philosophical Approach*, New York: Marcel Dekker Inc.

Marsden, D. (2003) 'Rights, Culture and Contested Modernities', paper presented at the INTRAC 5th Evaluation Conference, Kontakt der Kontinenten, the Netherlands, March 2003.

Morgan, P. (1999) 'An Update on the Performance Monitoring of Capacity Development Programs', paper presented at the meeting of the DAC Informal Meeting on Institutional and Capacity Development, Ottawa, May 1999 (originally prepared for CIDA's What Are We Learning series).

Neely, A. (2002) *Business Performance Measurement – Theory and Practice*, Cambridge: Cambridge University Press.

Neely, A. (2003) 'Gazing into the Crystal Ball: The Future of Performance Management', paper written for Centre for Business Performance, Cranfield School of Management.

Neely, A. and Adams, C. (2000) 'Perspectives on Performance: The Performance Prism', *Focus Magazine for the Performance Management Professional*, online publication, August, 4.

Neely, A. and Bourne, M. (2003) 'Why Measurement Initiatives Fail', paper written for Centre for Business Performance, Cranfield School of Management.

Oakley, P. (2001) *Evaluating Empowerment – Reviewing the Concept and Practice*, Oxford: INTRAC.

Oakley, P. Pratt, B. and Clayton, A. (1998) *Outcomes and Impact: Evaluating Change in Social Development*, Oxford: INTRAC.

Ospina, S., Diaz, W. and O'Sullivan, J. F. (2002) 'Negotiating Accountability: Managerial Lessons From Identity-Based Nonprofit Organisations', *Nonprofit and Voluntary Sector Quarterly* 31(1): 5–31.

Oster, S. M. (1995) *Strategic Management for Non Profit Organisations*, Oxford: Oxford University Press.

Paton, R. (1996) 'How Are Values Handled in Voluntary Agencies', in D. Billis and M. Harris, *Voluntary Agencies – Challenges of Organisation and Management*, London: Macmillan.

Paton, R. (2003) *Managing and Measuring Social Enterprises*, London: Sage.

Raine, W. J. and Willson, M. J. (1997) 'From Performance Measurement to Performance Enhancement: An Information System Case-Study from the Administration of Justice', *Public Money and Management* 17 (1): 19–25.

Rees, J. L. (1992) 'Managing Quality in Public Services', in C. Duncan (ed), *The Evolution of Public Management*, London: Macmillan.

Santos, S., Belton, V. and Howick S. (2001) 'Adding Value to Performance Measurement by Using System Dynamics and Multicriteria Analysis', Research Paper No.2001/19, Strathclyde Business School.

Sheldon, T. (1998) 'Promoting Health Care Quality: What Role Performance Indicators?' *Quality in Health Care* 7 (Suppl): S45–S50.

Smillie, I. and Hailey, J. (2001) *Managing for Change – Leadership, Strategy, and Management in Asian NGOs*, London: Earthscan.

Smith, P. (2000) 'A Comment on the Limitations of the Logical Framework Method, Reply to Gasper, and to Bell', *Public Administration and Development* 20 (5):439–41.

Wood, A., Apthorpe, P. and Borton, J. (2001) *Evaluating International Humanitarian Action – Reflections from Practitioners*, London: ALNAP and Zed Books.

www.alnap.org, visited April 2004.

Occasional Paper Series No 44

MEASURING SUCCESS

Issues in Performance Measurement

John Hailey and Mia Sorgenfrei
INTRAC Praxis Programme

NGOs are under pressure to invest in evaluation and impact assessment, as funding constraints, calls for accountability, and concerns about quality and effectiveness have led to donor demands for more sophisticated performance measurement strategies.

Performance measurement is a rapidly changing field, beset by methodological difficulties. Historically, measurement systems have undergone similar evolutions in the public, private and non-profit sectors – from product to process orientation; from quantitative to qualitative methods and indicators. New or modified frameworks are frequently put in place in attempts to address arising issues crucial to ensuring quality measurement. Such issues are particularly visible in the development and relief arena characterised by complexity, unpredictability and continuous change. Practitioners conducting evaluations face operational challenges such as how to choose the appropriate approaches; how to apply them in a culturally sensitive way and adjust measurement systems to the specific context; how to ensure stakeholder participation; and how to mobilise adequate resources for the evaluation.

ISBN: 1-897748-84-1



Postal address: PO Box 563, Oxford, OX2 6RZ
Registered and visiting address: 65 George Street, Oxford, OX1 2BE

Tel: +44 (0)1865 201851 Fax: +44 (0)1865 201852
Email: info@intrac.org Website: <http://www.intrac.org>

INTRAC is a company limited by guarantee No. 2663769
Registered in England A Registered Charity Charity No: 1016676