Can Accountability Systems Enhance Program Quality?

John Owen

John M Owen
j.owen@unimelb.edu.au
Centre for Program Evaluation
The University of Melbourne

Abstract

Accountability is a catch cry in today's public sector politics. It is one of the slogans used to lobby for more and better evaluation, and is seen as a good. How can any manager bearing responsibility say that she should not be accountable? This is due to a perception that funds allocated to education, health and welfare should be spent wisely.

But what does accountability mean in practice? How should evaluation be designed within an accountability framework? The recent emphasis on performance assessment has adopted a payment by results mentality. This involves 'reporting up' of outcomes to a higher authority with a view to justifying resources used for the results obtained.

However, our experiences suggest that such an approach to evaluation produces little helpful information to higher authorities.   All too often, these reports contain simplistic information which, in the end, is not brought to bear on decision making about the program or organisation under review.

Also, relationships built on conventional accountability are all too frequently adversarial in nature. This leads to situations where managers go to extreme lengths to justify what they have done, in order to protect themselves, their organisations and their budgets.

A more useful expansive view of information for accountability is required. This retains the provision of information about strategic outcomes, but also requires higher authorities to sanction processes that are likely to lead to organisational learning and policy improvement. That is, the fact that an organisation makes a commitment to organisational learning via evaluation should be recognised by a funding agency as a form of accountability.

This presentation will discuss a definition of accountability based upon these notions. We suggest that managers be held accountable for:
•setting up defensible internal evaluation procedures
•seeking advice and support from external evaluators,

•asking questions related to organisational problems
•ensuring that staff follow through on evaluation commitments, and
•use the findings to make organisational and program improvements.

Introduction

This paper focuses on meanings of accountability and the use of evaluation to enhanced the quality of programmatic interventions (policies and programs, hereafter programs) in social and educational systems. There is no doubt that accountability is currently a 'buzz' word in the private sector. The recent collapse of large corporations in the United States, Australia and elsewhere has led to an erosion of confidence in the quality of financial accounting standards and the behaviour of senior executives. The very fabric of economic life in western societies has been called into question. Governments are now introducing measures designed to restore confidence in the stock market, and to punish company directors and others who have attempted to defraud their shareholders.

Concern for accountability is also current in the public sector. There is evidence that the public needs to be satisfied that spending on public sector programs and institutions is both effective and efficient (Gilbert 1998). Accountability is a catch cry in today's public sector politics.

Associated with the increased emphasis on accountability is a connection with the use of systematic evidence to show that programs are effective and efficient (Ryan, 2002). The need for evidence is particularly evident in the United States following the passage of the GPRA Act in 1993, which requires federal public authorities to specify and report on the achievement or otherwise of stated program objectives (Gordon and Heinrich, 2004).

However, there is a danger that strident calls for accountability can lead to systemic arrangements which can have deleterious effects on the quality of programs and the clients of these programs, resulting in outcomes which are the opposite to those they are designed to achieve (Earl and Lafleur 1997). I take this issue up later in this discussion.

Meanings of Accountability.

What do we mean by accountability? Related to the education sector, accountability has been described as:

'the responsibility for the justification of expenditures, decisions, or results of one's own efforts. Program managers and teachers should be, it is often said, accountable for their salaries and expenditures and time, or accountable for pupils' achievement, or both' (Scriven, 1991; p46).

In terms of programmatic interventions an essential aspect of accountability is the obligation to report on the exercise of responsibilities conferred on program managers for the expenditure of funds and implementation of mandated activities.

Accountability thus seems to involve a transaction along the lines of 'if we give you the resources, we expect you to tell us what you have achieved with them'. For an individual professional practitioner the resources are in form of a salary, and for a programmatic intervention, resources that fund the intervention. The conventional accountability process transaction is based around the flow of funds in one direction and the flow of related information in the other.

How does accountability differ from responsibility? Well, we think that they are similar with the exception that accountabilty has greater consequences for the personnel involved in providing the intervention under review. Accountability implies the inclusion of sanctions whereas responsibility does not. Sanctions could come in many forms and are worn by those responsible for the intervention being delivered. The result of an accountability transaction might be a reprimand for the program director, a note on an employment file, a promotion denied, or even the loss of job.

I recall a senior public servant responsible for the radical changes in school governance in the early 1990' telling me that, if the findings of a school review showed that a school was 'failing' the school could be closed and/or a principal could be dismissed. This raises the issue of attribution and fairness. It seems that one can only be accountable for actions over which one has direct control, a point made by Lee Cronbach more that two decades ago. However, we are aware within the Westminster system of government that a Minister is responsible for her department, and that, if a major impropriety occurs, the Minister traditionally takes the blame for this indiscretion and, in sone cases, offers to resign. This is despite the fact that the Minister cannot possibly exert control over all the events that are handled by the department or the decisions of public servants at all levels within the bureaucracy.

Concerns about consequences are a key concern for senior staff responsible for implementing government policy. While they live with accountability on a day-to-day basis, these concerns are heightened when a review is undertaken of interventions for which they have a responsibility. In the Australian context these can be in the form of performance audits, and when the situation is seen to be more grave, a Royal Commission. Both of these reviews are evaluative in nature, with an emphasis on looking back at outcomes and their causes.

These examples give me the opportunity to segue into the next section of this paper, namely the relationship between accountability and evaluation.

Roles for Evaluation in Accountabilty.

As we have indicated earlier, a new age of evidence-based decision-making has emerged which has implications for accountability systems. Observers suggest that, while professional judgement and opinion were often regarded as sufficient information, they no longer have widespread acceptance as a means of confirming the effectiveness of programs (Reed and Brown, 2001).

This seems an appropriate time to summarise trends in these systems over the past decade. In doing so, we rely heavily on recent North American based research literature, and ask readers to extrapolate these findings to their own contexts.

Accountability systems which have emerged over the last decade are characterised by the following

*increased community interest in what has been achieved for the expenditure of funds on social programs (Segerholm, 2003), and more attention to the needs of consumers in the planning of public programs

*devolution of authority from the centre to departments and agencies with the proviso that they will report on program effectiveness

*evaluation being increasingly integrated into accountability arrangements at whole of government and departmental levels

*differing accountability imperatives depending on position in the program provision hierarchy. For example, (1) fiscal and overall performance reporting 'up' from government departments to centralised fiscal agencies responsible for budget allocations, and (2) emphasis on program performance within a department from program manager to the senior executive (Ryan, 2002)

*the likelihood that evidence based systems in organizations cannot meet the accountability needs of all stakeholder groups

*lack of evaluation expertise at department and agency level, particularly not for profit agencies, which impacts on the quality of evidence collected and disseminated (Poole et al 2000), and the need to divert program development resources to evaluation (Ryan et al, 1996)

*evaluation being incorrectly equated with the development and application of performance information and the inability of indicator systems to adequately tap the range of outcomes of a given intervention (Schalock and Bonham, 2003)

*widespread concern that accountability information which is reported up is not being used for decision making at the next level of the hierarchy (Leighton 2003)

*the inability of outcome focussed frameworks to provide practitioners with information will lead to improvements in program provision, despite this being a key rationale for embracing accountability systems (Ryan, 2002)

*a need for champions of accountability to embrace a wider view of what counts as evidence (Wholey, 1994, Wandersmann et al, 2000).

This summary provides a clear message that normative accountability systems have serious flaws. For example modal evaluative approaches rely heavily on the collection of performance indicators. The evidence being collected is partial and incomplete, it does not give a true picture of what is being achieved, and there is little indication that the findings are being used in any meaningful way by senior decision-makers. Under these circumstances, there is no a sound basis for making accountability decisions. It may not be surprising that we know of no senior manager who has been demoted or lost a job on the basis of any evaluative activity that has purported to be undertaken for accountability purposes.

One can think of these approaches being akin to the bottom line in business and industry. Show us the profit, don't worry about how you got there is a catch-cry in the financial pages. We see the accountability movement in the public sector as having been hijacked by economists, politicians and others who subscribe to the bottom line mentality, with the transfer of a methodological technology that fails to grasp the complexities of social interventions

Our view is that accountability is an important facet of the way democratic societies function. We also agree that the collection and use of evidence based on the use of social science methodologies is an important contributor to accountability related decisions. But what we take issue with is

(1) the narrow use of performance information methodology that has been used in accountability systems

(2) the excessive pre-occupation with measuring outcomes without regard to the complexities of the interventions which are supposed to produce them.

(3) the normative view which assumes that those at the top of a system hierarchy, such as senior policy makers, should be the primary beneficiaries of an accountability system, in terms of how it is structured and the nature of the information that is assembled.

In the following sections we suggest ways in which accountability systems can be more aligned with the need to provide quality interventions.

Expanding the Range of Evidence in Determination of Program Impact

Reliance on performance information and in particular the use of simple performance indicators has been criticized by several authors, the most outspoken on whom has been Perrin (1998). Chief among these criticisms is the claim that indicators give an atomistic rather than holistic representation of achievements, that they are reductionist, cannot measure unintended outcomes, tend to stress easy to measure aspects of the program, and fail to take into account unintended outcomes. Critics also claim that the use of specified indicators encourages goal displacement, the phenomenon of orienting program delivery towards the achievement of higher scores on the indicators, especially those that are valued by management.   There are also issues related to the aggregation or amalgamation of individual indicator scores and the development of target levels by which the performance of programs will be judged. In practice, these evaluation logic issues are rarely considered in the development of indicator systems.

It seems that more comprehensive set of data collection and analytical techniques, using a comprehensive mixed-method design provides a more valid and acceptable approach to determination of outcomes. The example of the work of the Auditor General in Victoria illustrates a trend in this direction over the past decade. The Auditor General is now committed to mixed method designs for all its performance audits in addition to an increased emphasis on the 'triple bottom line'. This is not to say that indicators are never used, in fact they are very important in these evaluative studies. However, their use in now subject to examination so that they are consistent with evaluation logic discussed above and combined with other methodologies, eg case studies of selected cases, in order to provide a more complete and rounded findings set.

Accountability and Quality Assurance

We believe that preoccupation with outcomes in accountability systems may in fact be a barrier to, rather than an assurance of quality. For us, accountability seems to be consistent with a quality assurance focus rather than an outcomes focus (Ciarlo, 1982)

In a quality assurance focus, accountability procedures need to be invoked as the beginning of the policy intervention process. Others agree with this position. For example Wholey (1994), Wandersman et al (2000), and Ryan (2002) view accountability in terms of evidence that shows that managers have attended to program integrity. Programs need to be based on the needs of clients, have taken into account the latest research on the problem and its solution. This implies the need for program managers to provide richer and more diverse information than outcome measures. In such a system accountability moves forward to the inception of a program and then requires evidence to be progressively provided as the program is implemented.

Accountabilty Up <u>and Down</u>

We believe that there are alternative ways for a multi-level social program delivery system to implement accountability procedures (Herrel and Straw, 2002) that are more even handed. By a multi-level system we mean a centrally developed policy/ big P Program which is ultimately delivered by providers at the local program level. This schema relies on the need for central policy makers, who could be regarded as the proponents or auspicers of the given Program, to be accountable DOWN to providers in addition for program providers to be accountable UP.

How would this more symmetrical approach to evaluation and program provision operate across a system?

For accountability DOWN the following elements need to be in place. The auspicing organisation would
•provide clear policy regarding the ways each provider in the system undertakes internal program development <u>and</u> evaluation.
•negotiate with all providers about what findings need to be reported UP to the auspicer
•provide the resources for providers to undertake meaningful and valid development and evaluation, and assistance to use the findings internally
•disseminate 'usable evidence' to providers. This could be in the form of indicators about the core business of the Program (big P), which are more efficiently collected by the auspicer than by individual providers, yet can be useful to individual providers (Earl and Katz, 2002).

For accountability UP, the following elements need to be in place. All providers in the system would
•report on key outcomes, which are of use to the auspicer for <u>their </u>accountability to funding purposes, for example to a government department of treasure or financial management
•show that they are using relevant evidence and findings for internal decision-making and program improvement.

This approach to evaluation relies heavily on the ability of providers to undertake meaningful internal evaluations which lead to organisational learning. We think of such a scheme as being also consistent with the 'quality assurance' approach, which was mentioned in the previous section. A quality assurance approach is based on auspicing organisations making it mandatory for the providers to incorporate evaluation into local level program provision.

A quality assurance framework is designed to serve the information needs, first and foremost, of the agencies delivering their own program, and thus its major audience is internal to the program, managers and program deliverers. Accountability to the auspicing organisation comes by reporting that internal evaluation has taken place and that it is contributing to the effectiveness of big P Program delivery on an ongoing basis.

In summary, a quality assurance scheme requires program providers to use evaluation information for all stages of program provision. For many providers this represents a sea-change in their approach to what they do and can be summarised by the following:

The use of a quality assurance scheme would mean that we would no longer think of providers only as knowledge distributors. Providers would also be centres of enquiry, which is consistent with organisational learning. They would be creators as well as distributors of knowledge. It has been said that a provider which is engaged in organisational learning is an agency characterised by a pervasive search for meaning in its work.'

Implications for Auspicers and Providers. Such a scheme cannot become routine without the support of central auspicing authorities. They need to be committed to an expanded and somewhat different notion of evaluation compared to one based on performance measurement.

There are two key ideas here.

The first is that reporting for accountability purposes involves indications that a provider agency
*complies with the the intentions of the big P program, and
*uses evaluation to assist internal decision-making, along the lines of what has just been described. That is, evidence of the use of evaluation in program decision-making must be supplied to the auspicer.

The second is that resources must be provided in program budgeting for internal evaluation. This also includes resources for training managers and other key staff about an approach to evaluation which will be new to many of them. There will also be a need for support for professional evaluators to work with program staff. I envisage that staff who work in these ways would have some knowledge about organisational learning and participatory principles which have been encouraged by Cousins, Patton, Torres and others (Cousins and Earl 1992), (Patton 1997), (Torres, Preskill et al. 1996).

Our experience with procedures such as these is that, over time, an agency which sees the benefits of undertaking internally useful evaluations builds an evaluation culture (Owen and McDonald 1999). This can lead to the desirable situation of creating a learning orientation among staff in the agency, an increase in professionalism and increased effectiveness and efficiency.

Concluding Comment

This paper makes the claim that, as they stand, normative accountability systems do not work. They appear to waste the time and effort of program managers, do not provide useful information to central policy makers and do not add to the quality of social and educational programs. They also represent a waste of fiscal and personnel resources.

Several suggestions to remedy this situation are put forward. We hope that these suggestions can begin a more fruitful discussion and debate about roles that evaluation and evaluators can play in the contributions of our discipline to accountability, quality programs and good government.

References

Ciarlo, JA. (1982). Accountability Revisited: the Arrival of Client Outcome Evaluation. Evaluation and Program Planning, 5, 1, 31-36.

Cousins, J. B. and L. M. Earl (1992). "The Case for Participatory Evaluation." Educational Evaluation and Policy Analysis **14**(4): 397-418.

Earl, L. and Katz (2002). Leading School in a Data Rich World. In Leithwood, K, Hallinger, P, Furnman, G, Gronn, P, MacBeath, J, Mulford, B and Riley, K (Eds.). The Second International Handbook of Educational Leadership and Administration. Dordrecht, NL: Kluwer (in press).

Earl, L. and C. E. Lafleur (1997). Forging a Link between Accountability and Improvement in Schools. Orbit **28**(3): 3-52.

Gilbert, A., D (1998). Ensuring Accountability. Melbourne, Australia, The University of Melbourne.

Gordon, RA and Heinrech, CJ. (2004). Modelling Trajectories in Social Program Outcomes for Performance Accountability. American Journal of Evaluation, 25, 2, 161-189.

Herrel, J., M and R. Straw, B (2002). "Conducting Multiple Site Evaluations in Real World Settings." New Directions for Evaluation **94** (Summer 2002).

Leighton, B. (2003). Evaluation and Accountability. Paper presented at the annual meeting of the American Evaluation Association, Reno, November 2003.

Owen, J. M. and D. McDonald, E (1999). "Creating an Evaluation Culture in International Development Cooperation Agencies." Journal of International Cooperation in Education **2**(2): 41-53.

Patton, M. Q. (1997). Utilization Focused Evaluation. Third Edition. Thousand Oaks, CA, Sage.

Perrin, B. (1998). "Effective Use and Misuse of Performance Measurement." American Journal of Evaluation **19**(3): 367-379.

Poole, DL etal. (2000). Evaluating Performance Measurement Systems in Non-Profit Agencies: the Program Evaluation Quality Scale. American Journal of Evaluation, 21, 1, 15-26.

Reed, CS and Brown, RE. (2001). Outcome-asset Impact Model: Linking Outcomes and Assets. Evaluation and Program Planning, 24, 3, 287-295.

Rist, R. (1995). "Postscript: Questions and Answers." New Directions for Evaluation **95** (Fall 1995): 167-174.

Ryan, K, Geissler, B, and Knell, S. (1996). Progress and Accountability in Family Literacy. <u>Evaluation and Program Planning</u>, 19, 3, 263-272.

Ryan, K. (2002). Shaping Educational Accountability Systems. <u>American Journal of Evaluation</u>, 23, 4, 453-468.

Schalock, RL and Bonham, GS. (2003). Measuring Outcomes and Managing for Resultys. <u>Evaluation and Program Planning</u>, 26, 3, 229-235.

Scriven, M. (1991). <u>Evaluation Thesaurus</u>. Newbury Park. CA, Sage Publications

Segerholm, C. (2003). Researching Evaluation in National (State) Politics and Administration. <u>American Journal of Evaluation</u>, 24, 3, 353-372.

Torres, R. T., H. S. Preskill, et al. (1996). <u>Evaluation Strategies for Communicating and Reporting: Enhancing Learning Organisations</u>. Thousand Oaks. CA, Sage.

Wandersman, A et al. (2000). Getting to Outcomes, A Results Based Approach to Accounability. Evaluation and Program Planning, 23, 3, 389-395.

Wholey, J. S., H. P. Hatry, et al. (1994). <u>Handbook of Practical Program Evaluation</u>. San Francisco, Jossey Bass.

Conf 04 acc pap final 070904ii