Performance Measurement in Perspective

An exploratory study of Australian foundations’ approaches to performance measurement

September 2014

ACPNS Working Paper no. 64

Professor Diana Leat

Alexandra Williamson

Dr Wendy Scaife
The Australian Centre for Philanthropy and Nonprofit Studies (ACPNS) is a specialist research and teaching unit within the Queensland University of Technology Business School in Brisbane, Australia.

It seeks to promote the understanding of philanthropy and nonprofit issues by drawing upon academics from many disciplines and working closely with nonprofit practitioners, intermediaries and government departments. The mission of the Centre is “to bring to the community the benefits of teaching, research, technology and service relevant to the philanthropic and nonprofit communities”, with a theme of “For the common good”.


CRICOS Code: 00213J
September 2014
# Table of Contents

Acknowledgements ................................................................................................................................. 5  
Five Key Findings .................................................................................................................................... 6  
Executive Summary .................................................................................................................................. 7  
Introduction ............................................................................................................................................. 11  
  Background........................................................................................................................................ 11  
  The Study ........................................................................................................................................... 11  
  The Literature ................................................................................................................................... 11  
Methods ................................................................................................................................................. 12  
Data and Findings ................................................................................................................................... 13  
  Descriptive Statistics .......................................................................................................................... 13  
  The Pressure to Measure in Australian Foundations ........................................................................ 14  
  Measuring the Foundation’s own Performance ................................................................................. 15  
  The Pressure to Measure in individual Foundations? ....................................................................... 16  
Sources of the Pressure .......................................................................................................................... 17  
  Application and Exceptions ................................................................................................................. 19  
  Approaches and Styles ......................................................................................................................... 19  
  Purposes of Measurement .................................................................................................................... 25  
  The Consequences of a Disappointing Report ................................................................................... 27  
  The Consequences of an Excellent Report .......................................................................................... 30  
Dissemination ......................................................................................................................................... 32  
  Resources for Measurement ................................................................................................................. 33  
  Techniques and Sources ....................................................................................................................... 33  
  Considering the Complications ........................................................................................................... 33  
Costs and Benefits .............................................................................................................................. 35  
  Risk and Measurement ......................................................................................................................... 36  
  The Future: More or Less? .................................................................................................................... 37  
Discussion and Conclusion .................................................................................................................... 38  
  Overview ............................................................................................................................................ 38  
  Some Dilemmas ................................................................................................................................. 39  
  Key Findings ...................................................................................................................................... 40  
Suggestions for Foundations .................................................................................................................. 41  
Limitations and Future Research ........................................................................................................... 45  
And Finally... ......................................................................................................................................... 45
Acknowledgements

In writing this report, we acknowledge with gratitude the thoughtful and honest contributions of the participants who gave up their time to be interviewed. We also thank the researchers and funders around the world who have undertaken and supported the previous work that informs our current knowledge.

Most importantly, we recognise the support of The Ian Potter Foundation, whose generous gift to Queensland University of Technology enabled Professor Diana Leat to spend three months at the Australian Centre for Philanthropy and Nonprofit Studies in late 2013, as part of the Centre’s Visiting Scholars program. We value both the gift and the understanding that lies behind it.
Five Key Findings

The study data was enormously rich and varied – thanks to the thoughtful ways in which respondents answered our questions. Here are five key findings (see page 40 for more detail):

1. The language of measurement is confusing and people in the sector use the same terms to mean very different things.

2. The link between ‘good’ and ‘disappointing’ performance and future funding is more complex and nuanced than many grantees might assume. Grant-makers do not typically see measurement in terms of praise and blame, and are very open to honest explanations of what went wrong and why.

3. Performance measurement has to be seen in the context of the grant-making process as a whole, not just tacked onto the end. Trust is embedded in the very nature of the grant-making process insofar as grant-making involves relying on more experienced others over whom grant-makers cannot exercise direct supervision and hold little authority.

4. Measuring has a number of different purposes which may be broadly categorised as: instrumental (e.g. to stop or continue something); conceptual (e.g. learning); and persuasive (influencing/satisfying donors or the board that money was well spent; involving new donors)

5. There was a seam of ‘rebellious realism’ running though a minority of interviews. These respondents accepted the movement towards measurement but treated the relevance of measurement with caution.
Executive Summary

Performance measurement in Australian philanthropic foundations is a hot topic. Foundation staff and board members are concerned with striking the right balance between their need for information with which to assess the effectiveness of their grant-making programs, and the costs in both time and money for grantees. Influenced by normative pressures, the increasing size and professionalism of the Australian philanthropic sector, and trends from the U.S.A. and the U.K., foundations are talking amongst themselves, seeking expert advice and training, consulting with grantees and trying different approaches.

Many resources examine methods of data collection, measurement or analysis. Our study instead treads into less charted but important territory: the motivations and values that are shaping the debate about performance measurement. In a series of 39 interviews with foundations from Queensland, New South Wales, Victoria and South Australia, we asked whether they felt under pressure to measure performance and if so, why. We queried whether everyone in the foundation shared the same views on the purposes of performance measurement; and the ways in which the act of performance measurement changed their grant-making, their attitude to risk, their relationship with grantees and their collaborations with other funders. Unsurprisingly, a very diverse set of approaches to performance measurement were revealed.

In general, Australia foundations saw ‘performance measurement’ as about monitoring (regular systematic collection and analysis of information to track the progress of program implementation against pre-set targets and objectives). Few engaged in evaluation (objective assessment of an ongoing or recently completed project, program or policy, its design, implementation and results) or impact measurement (the long term outcomes that may be attributed to the grant [and any associated activities] adjusted for what would have happened anyway, the actions of others, drop-off and displacement).

What came through strongly from each interview was the thoughtfulness with which Australian foundations are working on this issue. While actual performance measurement requirements for grantees varied from nil to multi-year external evaluations, in all cases the rationale was both considered and subject to review. Equally consistent was the understanding that while the principles underlying performance measurement stayed the same within a foundation, the application of those principles varied from grant to grant, changing with the size, term and purpose of the grant. Many foundations discussed and jointly agreed on performance measures with the grantee at the point of award of the grant.
Three interview questions in particular provoked strong and varied responses. They were:

- If the performance of a grant were disappointing/excellent, what would follow?

Contrary to initial expectations, a disappointing performance measurement of a grant did not lead to either termination of a staged grant (paid in instalments) or the loss of the possibility of future funding. Foundations almost uniformly responded that their expectation of a grantee following from a disappointing assessment would be an open and honest conversation about the reasons for the result, and a closer working partnership in future. In a similar reversal of the obvious, an excellent result from performance measurement of a grant would in fact reduce the chance of future funding from most foundations, on the understanding that the grantee was now well placed to seek funding from other sources, particularly public fundraising and government grants. Success led to congratulations and acknowledgement but also a sense of ‘they don’t need us any more’. These responses highlighted the importance of performance measurement to foundations as a teaching and learning tool for both themselves and their grantees: once you get it right, you move on.

- Do you feel that performance measurement has affected your appetite for risk?

Some of the foundations that responded to this question raised concerns about a focus on performance measurement inhibiting the willingness of Australian philanthropy as a whole to provide the ‘risk capital’ with which the sector is identified. These concerns were also reflected in comments on limitations of the scope of measurement in both theory and practice; and the definition of what ‘failure’ of a grant meant for a foundation.

- Do you measure your foundation’s own performance?

For a small number of foundations, indicators of their own performance included financial returns, operating costs, efficient and timely administrative procedures, and for some, fundraising targets. The engagement of family members was also mentioned as a goal for some foundations. The fit of a foundation’s grant-making within a larger context, such as a geographic community, a corporate social investment program, or an individual or family’s volunteering, was sometimes also considered. Internal assessment and review processes such as strategic planning days, surveying grantees, and benchmarking against other foundations were reported.

However, for the majority, measures of foundations’ performance were focused on the performance of their grantees – the foundation was seen as achieving its mission (or not) through the work of its grantees. This view saw the degree of success a foundation achieved as primarily
determined by its choice of grantee organisations (skill in choosing), and its contribution to the planning and direction of their work/project (adding oversight).

One respondent made a clear differentiation between what a foundation does, and what a foundation funds. This multi-layered approach considered the work of a foundation itself as building and developing relationships, undertaking research, making decisions (including saying no), providing gifts of money, and advocating and building networks. This distinction between the work of a foundation itself and the work of the organisations it funds was echoed in several responses, where a purpose such as engagement with staff, community or family existed for a foundation.

Australian foundations are feeling under pressure to measure the performance of their grants. They are questioning and exploring their own requirements, and are concerned about the onus they place on grantees. They are seeking to learn from each other’s experiences, from professional and industry training, from overseas foundations and research institutes, and from academia. Their motivations are to show that they are achieving results, to constantly develop and improve their grant-making, and to be broadly accountable to their stakeholders. Finally, performance measurement is a growing trend. All respondents see themselves doing either more, or importantly, doing it better, in a year’s time.

So, what does this study add to our existing knowledge of performance measurement in foundations? Much of the research in this field starts from the premise that measurement is both good and necessary, and that grant-making organisations that do not measure at minimum the outcomes of the grants they make are missing valuable information about their work. While not contesting this viewpoint, this study steps back to take a wider view of the landscape of performance measurement, to include the diverse range of approaches taken, the values and philosophies that underlie decisions around measurement, and the ways these decisions influence the ongoing work of the grant-makers and beneficiaries.

This working paper may interest grant-makers and grant-seekers of all types (policymakers), and the growing number of people in this country studying or reading about philanthropy. Individuals can access the paper from different angles, either by searching the contents page for particular points of interest or doing a keyword search of the document. The reference list may provide further reading for those wanting to know more. A literature review included as Appendix one offers particular benefit to students and reflective practitioners. Appendix two looks at various definitions of key terms from industry and academic sources and highlights some questions posed to judge the calibre of evidence provided.
The contribution of the Ian Potter Foundation in funding Professor Leat’s time in Australia, and the generous response from the many interviewees, have made this working paper possible. We thank them for their engagement with research that contributes not only to the growth of the Australian philanthropic sector, but also its effectiveness.
Introduction

Background

In the past decade emphasis on performance measurement has steadily grown across all sectors – public, business and nonprofit. In Australia, the government requested that the Productivity Commission research ‘the contributions of the not-for-profit sector with a focus on improving the measurement of its contributions’ (Productivity Commission, 2010, p. iv). Private foundations have not seen themselves as exempt from the pressure to measure. Previously largely content to do good; foundations are now choosing, or being strongly encouraged, to adopt aspects of ‘performance measurement’.

Much of the current debate and focus relates to techniques, measures and best practice, and a whole industry has grown up to respond to those concerns. Much less attention has been paid to other equally important questions about who and what is driving adoption of performance measurement, its meanings, purposes, costs, benefits and positive and negative unintended consequences.

The Study

The aim of this study was to encourage wider and deeper debate about the emphasis on performance measurement, and to document the approaches/methods used by Australian foundations and the background to those choices.

The study explored:

- How Australian foundations perceive the nature and sources of pressures for ‘performance measurement’
- Responses to those perceived pressures
- Perceptions of the meanings and purposes of ‘performance measurement’
- Performance measurement methods used, and any dilemmas
- Perceived costs and benefits of ‘performance measurement’
- Any perceived unintended, direct or indirect, consequences of the emphasis on performance measurement.

The Literature

The study began with a literature review focusing on previous similar research. While we found a mass of performance measurement advice and exhortation we found relatively little research of direct relevance. There were, however, some interesting findings from the literature review. One was that actual practice of performance measurement even among large U.S.A. foundations appears to be patchy. Another was that the uses of data from performance measurement are unclear. To read the full literature review, please see Appendix One (page 48).
Methods

On the basis of the literature review we constructed an interview schedule exploring the broad topics above. It was not feasible to use a random sampling technique so a purposive sampling approach was adopted to include a range of different types of foundation, of different size, age, geographical location and with different interests. Purposive sampling fits the exploratory nature of the study, as it allows for a range of forms and practices of performance measurement to be captured.

Initial recruitment was undertaken through the researchers’ professional networks in the philanthropic sector. Representatives from a range of foundations who have participated in previous QUT research (Scaife, Williamson, McDonald & Smyllie, 2012) and who indicated their willingness to be involved in further studies were also asked to participate. Further interviewees were recruited using a snowball sampling method, appropriate for this notoriously private and difficult to reach group (Biernacki & Waldorf, 1981).

In-depth interviews were used as the primary research method as they are “a highly efficient way to gather rich, empirical data” (Eisenhardt & Graebner, 2007, p. 28) with the intent of capturing the perspectives of staff, managers and board members of foundations. Interviews were undertaken in person where possible, or by telephone where not possible due to distance and cost barriers. The data collection for this initial exploratory study was conducted from October-December 2013 (see Appendix Three for a list of interview questions).

The researchers’ interview preparation included familiarity with each foundation’s website (where one exists) and any published reports. The interview protocol was piloted with colleagues who had previous experience in philanthropic grant-making. It included brief questions about the descriptive characteristics of the foundation, such as year founded, program areas, and value in dollars of annual distributions.

A QUT low-risk human ethics consent form was signed by participants before the interview commenced. The names of individual participants and the foundations they represented are not identified in this report for purposes of privacy and confidentiality.

Coding of interview transcript data was undertaken using NVivo. Use of software assists in coding, searching and linking data from different sources, particularly interview transcripts and documents (Marshall & Rossman, 2004, p. 181). Open or thematic coding was first undertaken to closely examine the data, categorising phenomena by theme and searching for patterns (Seale, 1999, p. 99) as an iterative process.
Data and Findings

Descriptive Statistics

Staff, CEOs and board members from a total of 39 foundations were interviewed. When grouped loosely by type, (and acknowledging the difficulties such classification entails), the sample comprised 6 community foundations, 17 family or independent foundations, 9 corporate foundations and 7 foundation management companies (Professional Services or Trustee companies). The average length of the interviews was 44 minutes. The interview transcripts were supplemented by review of annual reports and other foundation publications.

Figure 1: Sample of foundations interviewed by type, n=39.

The average age of the foundations was 28 years, although this figure was influenced by one charitable trust with an age of well over 100 years. The median age was 14 years.
The Pressure to Measure in Australian Foundations

We asked whether respondents believed that Australian foundations in general were under pressure to measure the results of their grants, and whether they felt under such pressure in their own organisation.

- The majority of respondents believe that Australian foundations are under pressure to measure the results of their grants. The pressure was generally seen as ‘high but not unbearably so’.
- Some questioned use of the word ‘pressure’.
- Those who did not believe that Australian foundations were under pressure to measure the results of their grants believed that this would come to be a pressure.

Differentiated pressure

According to some the pressure was differentiated both between types of givers and between types of institution.

‘There is a definite split in the clients that we work with between people who give in a heartfelt way, I suppose, and measure things quite emotionally and give in a much less rigorous way than they’d make a business decision .......It probably aligns slightly with those ones that sit around the family table and use the PAF [Private Ancillary Fund] as a tool to talk about, hey kids, you know? To get the kids involved. That’s part of the objective of the foundation, is to get kids involved rather than to save the world, to some extent’.

‘I feel the pressure is higher with professional trustee firms because organisations think - particularly nonprofit organisations - think that because we are upfront in getting a professional fee to help provide a philanthropic grant, that we should be more open than Bill who has his foundation with his local accountant and isn’t perceived to be paying a professional fee even though potentially there is’.

Similarly, some suggested that the pressure was greater for corporate foundations having to argue for their budget each year: ‘...warm and fluffy is the first to go. I need a way to prove that this is working for brand, employee involvement and so on’.
‘I wouldn’t say it’s pressure. I think there’s a desire to do it’

Lack of pressure

The minority who reported lack of pressure on Australian foundations usually linked this to what they saw as a wider lack of transparency in the sector.

‘There isn’t a great deal of transparency on the grants that are made. Sometimes that’s because the individual is so private. Sometimes it’s because the grant doesn’t go through any formal or public application process ... Because of the combination of those, then no, there’s not a lot of pressure on the foundations to evaluate’.

Measuring the Foundation’s own Performance

The majority of foundations made no attempt to measure their own performance and only a minority were, as they put it, ‘grappling’ with such measurement.

We also asked if foundations attempted to measure their own performance. Some respondents seemed to find the question unfamiliar. Some saw measuring the foundation’s performance as covered by assessment of grants: ‘I see that as one and the same, because the Foundation’s performance is a reflection of how well we distribute money to people who meet their outcomes’.

For the trust/foundation management companies the organisation’s performance was partly measured by its capacity to ‘spot winners’: ‘What do I want my program officers and managers to do? I want them to make grants within our themes that have the greatest impact that we can’.

Some foundations (especially those with living donors) pointed out that measuring a foundation’s performance would require clarity around criteria of success, and these might vary: ‘it’s not like they want to solve homelessness or solve a medical research dilemma, or whatever it might be. A real mark of success for them on their foundation is that their nieces, their nephews, their cousins, their children have been sitting on boards of whatever organisation they are passionate about, that they’ve been getting involved with fundraising, that they’ve been doing things, and that for them would be a mark of success of their foundation’.
There were also questions about how foundations might go about evaluating their own performance:

‘We do seek some feedback in the acquittal process, but we think that part is fundamentally flawed because we’re asking the grant recipient to evaluate a grant maker’. Another person noted: ‘The hardest part of all was coming up with our own program logic – what do we actually do?’

The Pressure to Measure in individual Foundations?

The majority of respondents felt under pressure to measure the results of grants in their own organisations. Again there was some quibble over the word ‘pressure’.

‘I think it is quite a strong pressure. I think it’s seen as one of the most important parts of my role’.

‘No I don’t feel under pressure because we do it. It’s an important part of our approach’.

Those who did not feel under such pressure in their own organisation explained this in terms of lack of Board/donor interest, different approaches to the purpose of giving and/or in terms of the perceived inappropriateness of measuring in giving.

‘It’s very much an emotional thing, this foundation, what they do. It’s what actually resonates with them as a family, and as individuals, as opposed to necessarily being about the end result’.

‘I think if you were to actually sit down and have a conversation with them around is measurement a good thing, they’d say, “Yes, of course.” They [the board] just don’t drive it’.

‘Philanthropy isn’t about business. It isn’t about measurement. It isn’t about onerous reporting requirements. Philanthropy is about love’.
And a caveat...

One person noted that measuring or not measuring could be of little consequence in practice: ‘often times they don’t even really read progress reports let alone actually try and delve into what difference it has done. The exception to that is if it is a pet project of one of the trustees...’

And another noted that family dynamics and ‘niceness’ could work against attempts to establish measurement: ‘everybody is just really nice. So everyone just wants to agree and yeah, no, we’re doing a good job, and of course we could do better, but it’s okay’.

Sources of the Pressure
We live in a world in which ‘measuring’ is increasingly important. We also live in what has been termed the ‘risk society’ – a world in which risks are everywhere, and nothing can be taken for granted. Given this wider environment all organisations, including foundations, might be expected to be under pressure to measure the results of what they do and assure donors, government and other stakeholders that their efforts are not merely well intentioned but actually produce benefits. On the other hand, foundations have often been seen as not subject to such pressures for a variety of reasons including the notion that this is ‘private’ money, that everything foundations do is ‘good’ almost by definition, and the more profound idea that fully endowed foundations survive irrespective of the success, or otherwise of their grants, and thus have no incentive or need to measure. So where do foundations see the pressure to measure coming from?

Respondents saw the sources of the pressure to measure as primarily of their own making. The pressure was seen as largely coming from within the philanthropic sector itself.

Evolution and imitation

‘I think as any sector develops and evolves, standards are raised and I think that’s what’s happened with philanthropy. I think it’s become a much more sophisticated, strategic sector’.

‘I think people look around and see what others are doing and they go “well gee whiz we’d better be in on this too”’.
Wider cultural values

‘It’s a values driven thing too of accountability, stewardship, inclusion, respect, truthfulness. If you try to live by those values you can’t just then be sloppy... or superficial in the way you evaluate. You’ve got to do things with a depth and a proper intent’.

New philanthropists

Some saw it as reflection of the styles of new philanthropists: ‘Some of it comes from the types of people that are getting involved in philanthropy now, who’ve been successful in the business world and want to use some similar sort of metrics to their philanthropy’.

But interestingly some of the community foundations suggested that measurement was not being driven by donors. ‘I think the flip side of it, again, is that not for profits are becoming more savvy. They’ve worked out they can engage people in a deeper way if they can talk about impact and outcomes in a better way.’

Online Information and Social Media

Some attributed changes to external factors such as the Internet: ‘With everyone now on Twitter and the way that we’re able to communicate and share information with our counterparts, particularly in the U.S. and the U.K., that’s what has driven a lot of it’.

Demands for transparency

Several people mentioned the effects of the wider culture of transparency and the need to be seen to be responsible: ‘I see pressure for transparency about what we’re doing, and how we do it. So that leads to a desire by us, if we’re going to have to be more transparent we want to be able to show that what we’re doing is worthwhile, which flips to a desire for performance measurement’.

Tax and self-regulation

Philanthropic donations attract tax benefits and some respondents mentioned this in relation to measurement. ‘The government is clearly aware of that [tax benefits to foundations] and clearly looking at it. It’s almost like self-regulation before someone comes and regulates us’.

Tight budgets

Only one person suggested that this was a resource driven movement: ‘It’s around getting the best performance from what is a scarce resource - whether it be dollars, or whether it be the skills of individuals that we can rally for our partners’.
Interest in measuring the foundation’s own performance (as noted above, not a widely shared concern) was seen as part of a wider culture of accountability and a fiduciary duty, and likely to increase.

Application and Exceptions

In general foundations were more concerned with assessing larger grants. Smaller grants were required to report but the format tended to be less detailed.

There was one exception to this generalisation. One respondent said that the foundation spent more time on assessing smaller grants because that was easier than assessing the larger ones.

Exceptions to measuring also included grants for tangible items and for very specialist work.

’We don’t spend much time on evaluating our very tangible grants – it’s to buy something or to provide something and they do it or they don’t. And at the other end of the scale how do you evaluate a piece of medical research?’

Approaches and Styles

’Impact’, ‘efficiency’, and ‘effectiveness’ were the terms most commonly used by respondents in talking about measurement, but as discussed further below these had a variety of meanings. There was little mention of Social Return on Investment (SROI) and if it was mentioned it was often in order to express scepticism.

Some respondents noted the difficulty that they could not impose too many conditions on the grant without risking the imposition of GST. Some also stressed the importance of pre-grant assessment and the way in which that could feed into post-grant assessment.
‘When an application comes through, if it’s not immediately clear in that application what the process of evaluating and measuring it is going to be, it won’t be successful’.

The general sense, however, was that as one person put it: ‘there’s no debate about whether we should do it. The only debate is how we do it, and of course there’ll be different points of view on that’.

A Continuum of Meanings of ‘Measurement’

Very broadly meanings of ‘measurement’ (Buckmaster, 1999; Hehenberger, Harling & Scholten, 2013; Maughan, 2012; Pritchard, Ogain & Lumley, 2012) vary along a continuum from:

1. **Reporting/monitoring** - This may be simple fiscal acquittal, and/or self-reporting at various stages against original objectives, use of key performance indicators etc.

2. **Stories** – a self-reported narrative account of the project and the ‘difference it made’.

3. **Evaluation** - If self-conducted then similar to (2), but more structured emphasis on including what worked and did not work, and why; may be formative (as you go along) or summative (end of project); may be self-evaluation or by an independent party. For definitions of evaluation, see Appendix two (page 59).

4. **Rationalist or technical measures** - For example, Social Return On Investment (SROI); outcome measurement against control, focused on results rather than process/or why something worked/happened.

5. **Impact** - the long term outcomes that may be attributed to the grant (and any associated activities) adjusted for what would have happened anyway (the ‘counterfactual’), the actions of others, drop-off and displacement (see for example, Hehenberger et al., 2013).

Figure 2. Types of performance measurement shown with increasing complexity from left to right

In general the further you go to the right along the continuum the greater the cost and time required, and (if done well) the greater the credibility of the evidence. It is worth noting that credibility is not always the same as utility or even reliability. We discuss ‘utility’ further in the Conclusion where we suggest that different approaches are adequate and useful for different purposes. Impact assessment is ‘the gold standard’ but is rarely seen as feasible in the foundation world for various reasons – it is also not useful for real time/immediate grant-making because it generally requires a 10 or 20 year time span.
So which types of measurement are most useful for the different purposes above?

<table>
<thead>
<tr>
<th>Type of measure</th>
<th>Most useful for</th>
<th>Less useful for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting/monitoring</td>
<td>Basic fiscal accountability of grantees and foundation managers</td>
<td>All other purposes; BUT if aggregated and analysed across a range of grants and then may be useful for other purposes including learning for future grant-making.</td>
</tr>
<tr>
<td>Stories and narrative reporting</td>
<td>Engaging family, employees etc. Basic, short term grantee accountability; basic short term foundation manager accountability; some use in learning for future grant-making.</td>
<td></td>
</tr>
<tr>
<td>Formative evaluation (as you go along)</td>
<td>Course correction; learning for future grant-making; learning for replication; some use in short term grantee accountability.</td>
<td>Enhancing reputation etc.; accountability of foundation managers; family etc. engagement Long term/high proof accountability of grantee and foundation managers</td>
</tr>
<tr>
<td>Summative evaluation (at the end of the project)</td>
<td>Grantee accountability; foundation manager accountability; some use in family etc. engagement; use in risk assessment, replication, enhancing reputation and learning for future grant-making when combined with other similar cases.</td>
<td>Long term/high proof accountability of grantee and foundation managers; replication; risk assessment.</td>
</tr>
<tr>
<td>Rationalist/technical</td>
<td>Accountability of grantees; accountability of foundation managers; some risk assessment; enhancing reputation and legitimacy (especially in business circles); learning for future grant-making.</td>
<td>Course correction Learning for replication Engaging family or other stakeholders.</td>
</tr>
<tr>
<td>Impact</td>
<td>Enhancing reputation and legitimacy of foundation in long term.</td>
<td>Any purpose requiring short/medium term data or results.</td>
</tr>
</tbody>
</table>
Respondents used a range of different terms to describe what they were doing, often using similar terms to mean different things.

The most common forms of measurement were at the left hand end of the continuum. Some respondents aspired to move closer to the middle (evaluation by independent ‘experts’).

Few expressed interest in rationalist/technical measures, and no-one expressed interest in impact measurement as defined here (the majority received their final report within months of the project end).

**Reporting/Acquittal/Monitoring**

‘So one aspect of the report is them to financially sign off and say that they’ve spent the money, or how much is left of it and why they haven’t spent it. That’s obviously very much a yes and no question type thing.’

Acquittal against previously, mutually agreed objectives was used by what appeared to be a growing number of foundations: ‘Our performance measurement is more around, well this was what you said in your submission you were hoping to do. Now you’ve used our money, have you done what you said you were going to do?’

Acquittal might also involve some level of reporting on what had been achieved: ‘There’s a measure of evaluation in the acquittal report. It’s probably just there so that the grantees feel that they pay some attention to measurement or evaluation, but it’s not a requirement. We want to know the result of whatever program we’ve funded, but in a very simplistic sort of way’.

Stories were also popular: ‘We actually also require a digital story, which some organisations found very onerous in terms of pulling that together. So it’s finding not only the information you include in your report, but also the tools you use in your acquittal process. But the stories are great. I know they found it hard but usually they’re great’.

Site visits were another addition to simple written acquittal reports. Visits had various functions: ‘The way we’re tending to do it is getting some of our Grants Committee going out and meeting with the grantees, and reviewing, and trying to get the facts, apart from the reports you get back. I think probably the value is in actually the committee members getting into the heads of the organisation...’
and really understanding what it’s doing. I think there’s a payoff there by understanding better what some of these organisations are trying to do and doing’. But visits were not an option for all because of time-poor trustees and staff.

Formal, detailed self or external evaluations appeared to be relatively rare although some foundations reported that they were now considering such methods.

What’s not working?

As discussed further below, respondents were generally keen to stress that they wanted grantees to feel that they could be honest. Some encouraged this by deliberately asking what was not working.

‘We ask that question - what didn’t work? What did you learn from that? How did you compensate? I think it’s just as important to find out and encourage a culture of a bit more honesty with funders, because if you don’t ask that question they won’t tell you what went wrong proactively’.

The limits of metrics

Only a handful of respondents mentioned numerical measures: ‘We also do another thing called Cost Per Participant, where we just look at how many people they’re supporting, what the level of funding is we’re giving them, and then divide it in a rough way to work out approximately how much the program costs per individual that is involved’.

But numerical methods were rare, and there was some anxiety about their value: ‘It’s neither appropriate nor effective or efficient to get hung up and overboard on metrics. The trick is to, I think, build as much robustness into your qualitative assessment processes’.

Measures and measuring should be grantee driven

This was a strong theme in a majority of interviews: ‘Dictating how they do it would end up being a box ticking exercise, and box ticking exercises aren’t worth it’.

‘Philosophically we don’t believe we should be acting as internal auditors on nonprofit organisations, because if you have that mindset you probably shouldn’t be funding the organisation. If you think you need to act as an internal auditor, you don’t trust the organisation and you shouldn’t be funding the organisation, or you should be investing in the organisation’s governance’.

‘It shouldn’t be the philanthropist’s job to second guess the organisation’.
Proportionality

Respondents were also generally keen to stress the importance of proportionality:

‘We don’t want to put the organisation through writing ‘War and Peace’ for a few thousand dollars’.

Approaches change

‘The more money they give away sometimes the more relaxed they become. People can be a bit overly anxious about their fears of what a charity is going to do with the money, and they get more confident as they get to know an organisation, and less worried in some ways about being too prescriptive about it. When people start off they want to know it’s for those 10 bricks on that wall. That’s what changes.’

Reporting can be a measure itself

‘We keep a database of all those score sheets, and then if we get a follow-up request from one of these organisations, that all goes into our database so that if we get a request from someone in the future, we’ll have a look through and say, “Look, the scorings of their reporting, there were a few issues” and we can look up and see they were a bit delayed in their reporting, or I had to really chase them to get their report, or the report really wasn’t specific to our project. It was too general, or whatever the case may be. It’s something that we’ve become a lot more rigorous around’.

Relationships and trust

What is appropriate in relationships between grant givers and grantees was an underlying theme in a number of interviews. For some the key issue was that a gift is not a gift if it is hedged about with conditions and reporting; for others the quibble was related to the complexity of the environment and the issues being worked on (see below for further discussion); for others the issue went to the heart of grant-making as fundamentally a form of devolved mission-pursuit based on trust.

‘All philanthropy is a trust investment in other people, which means I don’t even think it’s tasteful or ethical to try and put incredible pressure on a small agency to come up with some kind of incredibly exhaustive and prescriptive account of outcomes. I think it offends the principle of philanthropy. So it’s a fine line between offending the trust relationship and the investment in people. You have to trust that they’re not going to hoodwink you. You’re going to have to trust that if a project falls over in some part - the best outcome is if the trust relationship is there so they come back and say, “We’ve got a problem.” And we sit down and we discuss how to resolve it, which happens with us. Or they might come back and say, “We actually underspent by $1,700 and we were hoping that if we did this, this, this, that we can change the grant a bit to do that.” So we rely very, very strongly on upfront, honest relationships with our grant seekers and grant recipients. We work to their best sides, and they respond to our best side, I think’. 
Another person summed up the ideal funder-grantee relationship as ‘high trust, low docs, no surprises, and a learning environment’.

### Purposes of Measurement

Respondents reported various purposes of measuring the results of their grants including reputation, accountability, legitimacy, learning, personal satisfaction, and grantee learning and capacity building.

**Reputation, accountability and legitimacy**

‘Getting grantees to talk more about their performance, so you can tell those stories for fundraising purposes ... and, obviously, makes us accountable back to our stakeholders and supporters - that we’re granting to worthwhile projects that are actually making a difference as well’.

‘It’s how well are we doing in spotting the winners’.

‘It justifies our work. We need to make sure - we’ve got precious resources, both in terms of our time and our money - and we need to make sure that we are maintaining the vision and fulfilling the mission of the Foundation. Now the only real way we can do that is to measure the outcomes of our work against defined criteria. There’s no point us saying anecdotally, all these people seem happy in the picture. That’s not good enough’.

‘We’re a corporate foundation and for us it’s not entirely a branding exercise, but it is somewhat a branding exercise’.

**Accountability and learning for future grant-making**

‘We don’t mind sometimes if there is a failure, but you’d like to know why it is and then be able to avoid that next time around’.

‘To help us and the philanthropist understand the success of the grant, the learnings from the grant and the fit of that grant into other future benefits for the community still achievable’.

‘I guess the main thing we want to know is we put the money out for this purpose, is that being done? How effective has that been and would we want to do that again or not?’
Distinguishing between accountability and learning

‘Evaluation isn’t a yes or no for us - it’s a learning exercise. It’s about accepting that philanthropy is supporting new ideas and innovation. You have to accept that is not always going to work. Unless you know why it’s not working or why it is working, you can’t pass on the benefits to others and you can’t learn yourself. So, for us, it’s a tool about improving operations. For us, we wouldn’t use evaluation as a mechanism to say whether we would fulfil on an existing promise. All that’s going to do is get people to put together evaluations that are so weak that it cannot fail, and that’s not really helpful to anybody’.

Personal Satisfaction

‘... they just want to know that their money is having an impact - it’s sort of more often at that level than knowing that they’re actually shifting homelessness issues or something’.

‘They want to know the money has been well spent, on what they said it was, and they want to kind of be thanked for it and feel - feel good’.

And sometimes there is no purpose

One respondent described the frustration of assessing a grantee as performing very poorly year after year but this having no effect on decisions because the organisation was one of the elderly donor’s long term pet projects.

Grantee learning and capacity building

‘What we would like to see is that the information going into these progress reports is actually being used by them to evaluate their programs and adjust and report to their board, rather than be a one-off thing they do for us, because we think that if you leave those tools in place that will actually benefit them for wherever they get funding from, and also however they look to evaluate their programs. Part of achieving our mission is about creating internal capacity in a sustainable way for these organisations, and that’s also how to measure’.

Despite foundations’ hopes for positive benefits from performance measurement, some acknowledged that grantees might see it differently. ‘So they [the grantees] probably see it as a pain in the neck but necessary if they want to receive the money. That’s fair enough. I’ve run those organisations too but if you want to save yourself that problem, don’t accept the money.’

‘I think they sometimes see it as a bit of a spying exercise, looking over their shoulder, interfering in their work, and often they don’t see it as a useful tool for them to improve the performance of their
program or their organisation, simply because they don’t usually have time to think of it in those terms. It’s just a distraction from fulfilling their own mission.’

The Consequences of a Disappointing Report

It is probably true to say that many grantees assume that a poor performance report will lead to reduced or terminated funding, and an excellent report to further funding. In other words, measuring performance is about praise and blame, reward and punishment, expressed very directly in future grants.

- One of the key findings of this study is that foundations rarely use measurement data as a basis for punishment, and more often see it as a tool for course correction.
- Neither a disappointing nor an excellent result necessarily leads to less or more funding. The relationship between performance and further funding is more complex and nuanced.
- Reasons for terminating funding vary but a disappointing result is only one and not always a reason at all.

Disappointing reports are rare

The majority of respondents noted that disappointing results are rare: ‘It really doesn’t happen that often, believe me’. A few respondents mused on whether this was because they did not take enough risks.

Disappointing results are part of the territory

‘I think we’ve always taken the view that there’ll be some grants that we make that might end up being a dud’.

More fundamentally some respondents argued: ‘In the kind of work that we support, disappointing would have to be something fairly dire. We’re talking about homeless youth. The results of programs that are aimed at those are often what you could call disappointing’.

Some respondents linked ‘disappointment’ more specifically to the fact that they were often engaged in pilot funding: ‘We’ve had a couple like that but we also have funded quite a number of pilot programs, and so there’s going to be a higher possibility that they’re not going to be successful’.
The effects of a disappointing report

A disappointing report was rarely seen as a matter for punishment: ‘We don’t have a big ruler and a big stick and say, “Well you didn’t get there”.’

When asked about what would follow from a disappointing report, the most common response was ‘a conversation’: ‘So a discussion with the group to get more details obviously about what went wrong, but then lessons learned for us as well then for future rounds of grant making, looking at well, what signs should we have picked up before’.

Termination

The following approach was very rare: ‘We wouldn’t reinvest. Since our programs are about long term partnership, we won’t reinvest ….we would look at our grant process, our due diligence process, and say is there something in that we need to change as a result of this one not meeting its outcomes?’

This response was much more common: ‘No, we’ve never terminated a grant because of a poor report’. But several people noted that this was also related to the size and (short term) duration of grants.

Adjusting

Adjusting was by far the most common response to a disappointing report: ‘We need to adjust. We regularly adjust. If you apply for something and a priority is to achieve A, but somebody else comes in and achieves A three months later so you no longer need the funding, but you need B, we adjust to that. That is really common’.

Explaining responses

For a minority the lack of ‘blame’ had to do with the very difficult nature of the work undertaken. More commonly there was an acceptance that: ‘You can’t get the money back unless it is unspent’. ‘If they’ve spent the money I think we’ve got to write it off and say, “Well, that’s one we’ve got to chalk down to learning.” I mean we’re not going to get the money back and we’re not going to sue them’.

Another common reason for the rarity of termination was the importance the foundation or donor attached to the commitment given: ‘The family is very conscious when they’re making multi-year commitments. They are always subject to satisfactory reporting - there’s that line in it - but it would take an absolute disaster for them to pull their funding. In their mind, once a commitment is made they really need to do everything they can to honour that commitment’.
Less money or more?
Few respondents said that a disappointing result would necessarily lead to less money for the grantee, although this could happen in relation to pilot projects: ‘I can think of where we’ve done pilot studies and the pilot study hasn’t gone very well and they’ve come back to us and said, would you like to - there’s been an expectation that we’d fund the project, but where we’ve been unhappy with a pilot study we won’t proceed’.

As noted above, adjustment was a more likely response and in some cases might involve more resources, not less. ‘Putting in extra money because something is not working is very much a case-by-case thing. We would have to know why it’s not working, why they think it’s not working, and whether extra funding is needed or whether it’s extra staff that’s needed - why it’s not working for them to make some judgement and whether extra funding would seem to fix the problem. It might not always be the case’.

Relationships and due diligence
One strong theme in this section was the importance of relationships and the prior due diligence process. For some foundations assessment of performance is part and parcel of the wider funder-grantee relationship which is built during the pre-grant assessment process.

‘The thing is I don’t think it would ever get to the point where we would see it first in a report, because I make a point of catching up in person with the CEO, or at least the primary person who is responsible for delivering on the grant, three times a year, to get an informal verbal update, and we, generally speaking, get a read in those meetings if things are kind of going off track’.

The importance of scale
For some respondents assessing performance was an ongoing process based on close relationships and regular contact, but this was acknowledged to be much easier if the foundation had a smaller rather than larger number of grantees.

Honesty for learning
Honesty was a recurring theme in interviews. Funders wanted grantees to be open and honest about what had worked and not worked. Without honesty there could be no learning: ‘As long as people are honest with us and you can see that they’ve tried but it just hasn’t worked, well hopefully they’ve learnt something and perhaps we’ve learnt something’.

Being open and honest about things going wrong was the best way to overcome any funder disappointment with results: ‘The more open people are about it, the more likely we are to say, look,
happy to take another enquiry from you next year if you’ve got another project that you’re looking at. If you keep that communication open’.

Finally, whatever the content of the report some respondents suggested that grantees should receive feedback. ‘I reckon if you’ve put all that work into it and we go, “Oh, isn’t that good,” without formally acknowledging it - if I was the person writing the report I’d be really unhappy with that’.

The Consequences of an Excellent Report

The consequences of an excellent report were also complex. Continuation and termination of funding were both possible responses.

Further funding – or not?

In some cases the response was: ‘We’d give them more money, and we do that. In fact I’ve done it in the last week. Alternatively if it’s highly successful, if it’s shown to be highly beneficial, we might push more money into that program area’.

But another strong theme was to cease funding at the end of the existing commitment: ‘We say upfront to an organisation that we’re not here to fund you year in, year out ....but if it’s so good, or if it works so well, well then they’ve got to work out a way to put it into their regular program stream but they’re not going to have us funding it for the next 10 years. We try not to be locked into any one thing for years and years’.

‘You don’t get a bonus, if that’s what you’re asking’.

‘It doesn’t mean they’ll get more money. It just means that we’ll talk about it, and we’ll talk about it with our friends.’

In other cases the response was more qualified: ‘It would depend on where we are in the relationship. If the grantees is doing very well funding wise and the program is going fantastically, it would either mean that we continue to support it for a smaller period, although then we might look at well, now this organisation is doing well on its own it doesn’t need our support, so perhaps there’s another organisation out there that we could look at funding instead’.

More subtly, a good report might mean an increase in confidence in the grantee: ‘I don’t think it automatically translates into more money. What it might mean is we have a lot of faith in the people,
and if they come with something different to us we’ll look on that favourably.’

‘We would be talking about reinvestment and other ways that we can support them in non-financial support, so deepening the relationship with them, recognising that we will have a longer partnership with them. It also facilitates the funding discussions. It makes it easier to engage because we recognise that they can deliver’.

In other words, excellence may facilitate but does not guarantee future funding.

Promotion
In some cases the result of an excellent performance was simply a hope on the funder’s part that the project would be promoted: ‘What I’d love to see is they’re out there spruiking it to everyone else. You know, what a great grant it is’.

‘We’d celebrate that. We’d talk about it. We talk about it being a cracker. You know, it’s a cracker project’.

A minority of respondents took a more active role in promotion: ‘We will certainly story-tell about excellent pieces of work with a view to inspiring others in the sector. We will promote excellent grants. We’ve got on our web site great grant stories... We’ll have great grant outcome stories in each newsletter and so on ’.

‘I think we’d all sit around and sort of feel how clever we were, but then what we have to say is well, “How are we going to share this knowledge so that others can learn from it?”’.

‘If it’s really, really excellent we would think about talking with that organisation about maybe a media release, about how we can - in football terms - how we can handball it into places where people are more positioned to take action as a result’.

And for some types of organisation a good report was an opportunity to promote the organisation itself: ‘It will obviously get shared at the Board level, but that’s then part of the story that we promote and that we talk about more, and I guess if it’s a really good success story then that’s the one we focus on in general conversations we keep bringing up and hammering home the success of it’.

Responses may vary within one organisation
One respondent noted an important caveat in this discussion: ‘There was one (project) where it was quite easy for this organisation to pull this report together and say how many tonnes of x and y ... and blah, blah, blah, and they (the Board) read that and just went, “Wow. That’s amazing. Look at what we’ve done.” But they quite as easily wouldn’t have really been bothered if they didn’t receive that report’.
Dissemination

In general funders did not see themselves as responsible for disseminating the results of a project or grant. Some saw it as simply not their job or not appropriate because of issues of ownership, fairness, lack of demand and/or because they did not see themselves as advocates.

‘We’d expect them to do it. I think it just would not be fair for us to do that. We’re not there to judge organisations. We’re there to help them. They know more about their own business than we ever will. What we would hope is they would share the wins and also share the learnings and the negative learnings with others’.

Ownership was one obstacle: ‘We don’t disseminate the results of the evaluation, other than back to the grantee. In terms of external evaluations, well they’d own it so it would be their decision’.

Lack of demand was another: ‘I think if we were asked, for a good reason, to share our results, we would have no hesitation, but we’re not asked all that often’.

Unwillingness to play an advocacy role was another explanation: ‘They certainly wouldn’t see themselves as advocates. If an organisation has done something really well, been able to measure it, and then has gone out and managed to secure government funding or have gone out and changed a policy or whatever it is, they think it’s fantastic, but they don’t push it and don’t require it’.

The exceptions

Nevertheless some did sometimes see dissemination of results as part of their role and as a way of adding value over and above the grant: ‘Normally it would be the grantee, but there have been occasions where we will look at the results and actually there have been occasions where we’ve actually disseminated - with their approval - disseminated or brought organisations together if we’ve seen there’s a link between what you’re doing and what someone else is doing and brought them together’.

‘We often will host events for our grant partners around the launch or release of something or announcing something. We do a lot of that and because we’re a corporate foundation we’ve got that capacity to do that because we’ve got a building, we’ve got consumers, we’ve got caterers, so we can do that. So, yes, I’d say we, generally speaking, do try to’.
'So every 18 months what we try and do is get all the partners together for a day, or a day and a half, and the sub-text there is, “Why don’t we talk about what’s working and what’s not working”, in a community of learners’.

But sometimes lack of resources made this difficult: ‘it’s a medium term thing to have more the leadership role with the community groups to disseminate that type of information. Saying, “Have you thought about this?” or, “This is a story we want to share from this perspective.”... that would be an ideal situation to be able to have time and resources to do that’.

Resources for Measurement

In general responsibility within a foundation for overseeing measurement lay with the chief executive or senior program manager or was shared between grants officers.

In one organisation an extra member of staff had recently been recruited to allow the chief executive to spend more time on post-grant assessment. Time and resources were widely mentioned as a cost of engaging in measurement but there was little explicit mention of cost as a barrier to such activity. However, one person said: ‘We haven’t broken out how much we spend in measuring, because, I guess what would that tell us? How useful would that be? It would only frighten the horses’.

Techniques and Sources

Respondents’ sources of knowledge regarding measurement techniques were researching and reading on the Internet, their own professional experience, and from peers.

‘We just did different online research really to see how others were measuring performance.’

‘We’ve done quite a bit of sector consultation and understanding of other funders’

Considering the Complications

We asked respondents whether they tried to take into account the effects of external factors on results of grants, and whether they considered what might have happened anyway (i.e. without the grant).
Some clearly found these questions overly academic, but in general respondents were keen to emphasise that they considered the effects of external factors on the work of grantees.

‘People leave or get sick, government changes its policy, things don’t go as planned’.

‘Things go wrong. That’s life’.

‘It always evolves. If a project ends up the same way it started then it was probably a mistake’.

- A small minority of respondents said that they considered what would have happened anyway.
- The perceived lack of understanding of time and complexity in the social sector was more important to many than ‘what would have happened without the grant’.

20 year horizons

‘One of the problems with the business thinking often comes with people that are new to philanthropy and the not for profit world ... there’s so much more long-term, complex leadership challenges going on in the social sector that you have to educate people to why you have to look at 20 year time horizons, not 18 month time horizons and things’.

Success and failure are relative

‘I mean with homeless young people putting a roof over their heads might seem like it’s a 100% success, but it’s not because there’s so much more. They have addiction problems, they have personal problems, family problems, educational problems. So you can’t really say that giving them somewhere to sleep that’s not on the street is a successful outcome’.
Costs and Benefits

- The majority of respondents saw the costs of performance assessment for grant-makers as time and resources. The benefits were generally seen as learning and change.
- Costs to grantees were also seen as time and resources; some also saw costs in terms of diversion. Benefits to grantees were also seen as reflection and learning.

Learning and Change for Grant-makers and Grantees

‘So often we see philanthropists think they’re going to make their greatest difference doing A and they become quite disenchanted because A doesn’t achieve what they think it will, and they end up doing D, which they never thought they’d expect to do and getting immense satisfaction and seeing change’.

‘We don’t allow them [the grantees] to get their next payment until they submit their report. So they’re pretty good about getting the report to us. The benefits are that it instils a discipline in us. It instils a discipline in them’.

Confidence and A New Way of Thinking

‘I think it reassures us and the Board that we’re on the right track. It gives us confidence’.

‘I think it set the organisation on a whole different course. They (the grantee) weren’t able to get government funding beforehand, now they have been able to. Just that whole mind shift of no, we really need to be able to track what we’re doing with these kids, rather than just the feel good stuff’.

‘I think it has been incredibly useful for them’.

Future Fundraising

‘What I say to them [grantees] is, “Even if you don’t support this wholeheartedly, how much easier is it going to be to get the next grant if you can go in there and say, ‘This is what we measured. This is how we did it, and this is what happened at the end of the day.’” Because you just know that will resonate with a funder, so it’s in your best interest even if you don’t believe in it. That tends to get the cynics across the line’.
Costs for grantees

‘For the grantee - there are definitely negatives that can occur. If they’ve got a funder that’s wanting over-reporting on stuff that is not actually central to their mission but is central to that foundation’s mission, that’s a complete pain in the butt for the not for profits’.

‘My deepest concern is that we load our grantees up with all sorts of onerous time wasting, frustrating requirements because we can’t make our own nimble decisions about what’s happened, you know? Tell me in words of one syllable - what did you do? That’s insulting. If we funded it, we should know a lot more about what they did, and to take it back and say, “What did you do? We’re completely amnesiac and we’ve just got money.” That’s wrong. I think that’s bleeding our whole sector of time and effort to educate us and we don’t behave as if we should be educated. We behave as if we know everything already, so asking a clever question should be our first requirement’.

Caveats

Again there were caveats about benefits and costs. One person said: ‘often times they [the Board] don’t even read progress reports let alone actually try and delve into what difference it has done’.

Risk and Measurement

In general respondents believed that measuring made them willing to take more rather than less risk.

‘It enables us to take more risk because we can measure and evaluate more effectively and we can learn from previous challenges’.

‘I hope if evaluation is done well, and in conversation and in an educational way with the donor, it may increase their appetite for uncertainty and to realise that social issues change, and social programs are impacted by a whole range of things and a whole range of adaptive issues that don’t necessarily happen in the business world.’

But there were a few exceptions: ‘Yes I do worry that the Board could start looking at things and saying “oh they didn’t quite reach the target so we won’t fund that again’.”
Respondents did not generally believe that measurement had any effects on their program priorities.

However, two people said: ‘If we knew what was working it might mean we would do more of it’. Another suggested that more measurement might lead to more focus and ‘more multi-year funding arrangements and deeper engagement with a cause’.

The Future: More or Less?

- Respondents generally believed that they would be doing much the same or slightly more measurement in a year’s time.
- Doing it better was also a theme.

‘We won’t be doing less. I wouldn’t say we’d necessarily be doing more. We just hope to be doing it better’.

‘We may be doing less, or doing it differently because I want us to only ask for information we are going to use.’
Discussion and Conclusion

This discussion and conclusion draws together some of our key findings, dilemmas and reflections, and ends with some suggestions.

Overview

In the past doing ‘good’ was good enough in the foundation world. Today we live in societies where we see risk everywhere, nothing can be taken for granted, scandals about even the most trusted abound, and audit is viewed as an essential part of life (Beck, 1999; Power, 1994 and 2005).

‘Measuring’ is gaining ground in many institutions including the nonprofit sector where previously their very nonprofit-ness was read as a proxy for evidence of ‘good’. The rise of measuring in the nonprofit sector in the current cultural and financial climate is seemingly easy to explain: in addition to the wider culture noted above, resources are scarce, and the nonprofit fundraising market is highly competitive. If you want grants in a culture in which donors want the greatest bang for their buck then you have to produce ‘evidence’. Furthermore, if those around you – for example, other nonprofits, government funders – are providing or requiring evidence than you will be subject to isomorphic pressures (i.e. to conform to ‘best practice’ and/or the behaviour of peers) in order to meet your resource needs.

But endowed foundations are NOT resource dependent – they have resources and they, both in theory and in the past in practice, do not need to produce ‘evidence’ to secure existing or more resources. For endowed foundations nothing tangible happens if none of their grants ‘work’, and nothing happens if they are all huge successes – the foundation still survives (unless of course their investment policy is wrong). Community foundations and some corporate foundations and Public Ancillary Funds are in a different position and may feel under pressure to produce ‘evidence’ to legitimate expenditure and to encourage continuing donations.
Some Dilemmas

Arguably, for all types of foundation there are some dilemmas in measuring.

- There are legal dilemmas – what conditions can you attach to a grant before it becomes a contract and thus subject to tax?
- There are practical dilemmas – how do you measure and when do you measure? To measure impact (in the strict meaning) would require a 10 or 20 year time horizon.
- There are scale dilemmas – is it feasible to measure the results of 300 grants? And is it sensible to measure when the measuring costs more than the grant?
- There are philosophical/ethical dilemmas – is it appropriate to attempt to measure the effects of a gift?
- There are other cultural dilemmas especially in relation to the costs of measuring – foundations and their donors generally prefer to keep costs low.
- There are timing dilemmas – is it sensible to invest in measuring after the money has been spent (and when results will probably not be available in time for the next grant decision)?
- There are dilemmas with what counts as ‘success’ and ‘failure’. If a project achieves 50% of what it hoped to achieve is that success or failure?
- There are difficult issues to do with the purpose of the grant – is it to deliver a particular outcome, or to signal support? Or to build the capacity of a field or an organisation?
- There are difficult issues to do with the purpose of the foundation – is it about delivering social outcomes/‘changing the world’? Or about giving the donor or the company enhanced reputation? Or giving the donor(s) satisfaction? Or about involving the family? And so on.
- There are dilemmas with the uses of measurement information. Will the information be useable, and will it be used? If so, for what? And by whom?
- Grants are often only a part of the economic and environmental forces impacting the beneficiary. What are the conclusions that may be drawn from the measurement data? Will the information tell you who or what made the difference?
- Will measurement information be used in future decision making? Will it be of the right type and quality and available at the right time? And, in any case, are decisions based on factors other than hard data?
- Will measurement provide data that can be ‘used again’ in another setting and what will be its shelf life?
So given these dilemmas and, for endowed foundations the fundamental lack of necessity to measure, it is surprising in many ways that it is done at all. As one respondent put it: ‘while we set goals it’s not imperative that we meet them because no-one’s business is going to fall over because of it. It’s not like their own business when obviously they’re striving to meet goals because they’re accountable to shareholders direct. There aren’t those shareholders who are going to sack you all or anything like that if you don’t meet these goals’.

And yet, our study suggests that Australian foundations are talking about and working on measurement and generally see themselves as doing more in the coming years.

**Key Findings**

The data was enormously rich and varied – thanks to the thoughtful ways in which respondents answered our questions. Here we single out just five key findings:

1. The language of measurement is confusing and people in the sector use the same terms to mean very different things.

2. The link between ‘good’ and ‘disappointing’ performance and future funding is more complex and nuanced than many grantees might assume. Grant-makers do not typically see measurement in terms of praise and blame, and are very open to honest explanations of what went wrong and why.

3. Performance measurement has to be seen in the context of the grant-making process as a whole, not just tacked onto the end. Trust is embedded in the very nature of the grant-making process insofar as grant-making involves relying on more experienced others over whom grant-makers cannot exercise direct supervision and hold little authority. The pre-grant assessment process is the first building block of trust onto which more trust is built through an ongoing, more or less close and open, relationship between funder and grantee. For some foundations there was a clear feeling that if you did not or could not trust grantees then you should either not be a grant-maker or you should not be making grants to these grantees.

Scale of grant-making was an important issue here with smaller grant-makers celebrating the luxury of being in close, ongoing contact with every grantee, and large grant-makers accepting that size was in some respects a constraint on how they could operate.

4. Measuring has a number of different purposes which may be broadly categorised as: instrumental (e.g. to stop or continue something); conceptual (e.g. learning); and persuasive (influencing/satisfying donors or the board that money was well spent; involving new donors etc.)
5. There was a seam of ‘rebellious realism’ running through a minority of interviews. These respondents accepted the movement for measurement but treated the relevance of measurement with caution. They accepted that grantees had an interest in presenting a ‘pretty picture’, but more importantly they emphasised that a grant usually touched only a fraction of the problem/issue for a relatively short time. Respondents were aware that the most intractable problems are complex and multi-faceted and when you change one part you have only changed one part. Often a 20 year measurement horizon might be more appropriate than a one year horizon, and what counts as ‘disappointing’ has to be put in context and into a time-frame.

Suggestions for Foundations

The following are some discussion topics foundations may find useful.

- Clarity around the roles and purposes of your foundation and its grants
- Clarity and agreement regarding terms/language.
- Horses for courses: measures and purposes
- Encourage and reward honesty
- Re-consider the foundation’s role in dissemination
- Beware the tail wagging the dog
- More data are little help unless organised and used
- Beware ‘outcome creep’
- Puzzles and mysteries

Each of these topics is discussed in greater depth in the following pages.

Clarity around the roles and purposes of your foundation and its grants

A strategy for measurement requires clarification of the roles and purposes of your foundation. Is your purpose to change the world, or to bring a family together, or to involve employees, or to develop your business? Do you see yourselves as existing to fund what works? Or what doesn’t – yet? Not all grant-making is outcome driven: ‘The small grants program you might describe as being one more grounded in compassion than a very structured sense of trying to achieve outcomes other than what I call EBOs - evidence of the bleedin’ obvious. If a family hasn’t got a fridge you don’t need much of an outcome report when you give them one’. (The measurement purist would undoubtedly argue that if you want to measure the outcome then you need to measure whether the fridge improved the family’s diet and long term well-being, and the children’s performance in school etc. But barring the possibility that the fridge gets filled with soft drinks, chocolate and cream cakes, it seems reasonable
to suppose that in Australia having a fridge allows a healthier and more balanced diet than not having one).

**Clarity and agreement regarding terms/language.**

One of the most striking, but hardly surprising, findings was the ways in which the same terms are used with very different meanings (for a very brief review of the variety found in the literature see Appendix One). We suggest that it may be useful to distinguish between the following:

<table>
<thead>
<tr>
<th>Monitoring – often described as acquittal reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>This may be straight fiscal monitoring/accounting; and/or monitoring against previously agreed objectives and key performance indicators</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluation – understanding and assessing what worked, or did not, how and why.</th>
</tr>
</thead>
<tbody>
<tr>
<td>This may be formative – ‘a peek under the bonnet’; or summative – an ‘end of term’ assessment. It may be formal - undertaken by an independent specialist; or informal story telling by the grantee (purists might put story telling into another box)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rationalist/technical output or outcome measures:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using financial or other metrics, and often control groups, to measure outputs (immediate effects) or outcomes (wider/medium term effects) e.g. SROI; measures used in Social Impact Bond calculations etc. See, for example, <a href="http://www.gov.uk/social-impact-bonds">www.gov.uk/social-impact-bonds</a>.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Impact measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measuring the long term outcomes that may be attributed to the grant (and any associated activities) adjusted for what would have happened anyway, the actions of others, drop-off and displacement (see for example, Hehenberger et al., 2013)</td>
</tr>
</tbody>
</table>
The key difference between evaluation and much outcome measurement is that evaluation explores how and why something worked whereas outcome measurement is primarily interested in whether it worked. Impact measurement takes outcome measurement and evaluation a stage further and sets a very high bar. In this strict sense, we did not encounter any foundation attempting it.

As a general rule the cost and time required increase as you move up the scale from monitoring to impact measurement. Some would also argue that the reliability of the evidence also decreases as you move down the scale – but reliability may be different from utility.

When we asked respondents about the purposes of measuring very broadly they gave the following answers:

- Accountability of grantees
- Accountability of the foundation managers
- Course correction
- Risk assessment
- Learning for future grant-making
- Learning for replication
- Enhancing reputation and legitimacy (e.g. for raising funds/business)
- Engaging family, employees, community etc.

Horses for courses: measures and purposes

So the message is – decide what you want measurement for and then choose the least costly and most timely method for your purposes. For example, do not engage in costly long-term impact measurement if your primary goal is to ensure basic accountability of grantees and foundation managers, and/or to engage family members and enhance immediate reputation.

For a number of our respondents monitoring plus stories fulfilled their primary purposes of basic grantee and foundation manager accountability and family or employee engagement. It is difficult to learn very much from one grant – there are too many other factors at play – but looking at a whole program or a collection of similar grants in different places at different times may yield richer results. A few foundations were using this approach to identify recurring issues worth investigating further.
Encourage and reward honesty

Respondents emphasised how much they wanted grantees to feel able to be honest with them, and to tell them when things were going wrong so that the problem could be resolved if at all possible. Do you give your grantees ‘permission’ (e.g. explicitly asking what went wrong as though that were a natural expectation) and opportunities to talk about things that are not going so well on a regular basis?

Re-consider the foundation’s role in dissemination

Relatively few foundations saw it as their responsibility to play a role in disseminating results of projects that went well (clearly disseminating poor results is a more complicated issue). And some gave considered reasons why they did not play this role. It may be useful for every foundation to think through its position on dissemination, and whether this may sometimes be a way in which they could ‘supercharge’ the value of their grants.

Beware the tail wagging the dog

Are you ever attracted to program areas or projects because they are easy to measure and/or produce tangible results?

‘They do a lot of employment generation type work trying to get unemployed people into paid employment ... I don’t know that any of the family is particularly passionate about that, with the exception of perhaps the Chairman, but because it is so easy to measure and it’s easy to see results, it’s considered incredibly successful’.

More data are little help unless organised and used

Do you in fact use the data you already have? ‘We’re on a learning journey and as we develop we’ll be better at reporting processes that feed the information in ways that we want. The fact is we’ve got masses of data here but it’s not very well organised’.

Beware ‘outcome creep’

One of the often suggested risks of measurement in grant-making is that it will encourage funders to go for success and/or the ‘low hanging fruit’. Interestingly the majority of our respondents did not perceive this as a risk; indeed, many suggested that measurement/monitoring made them more willing to take risks because they were aware of the challenges ahead and could course-correct if necessary.
This raises the interesting possibility that the main value of measuring is not what comes out at the end but the thinking and analysis that goes in at the beginning; an idea summed up by DiMaggio (2002) in the suggestion that outcome measurement in the social world is probably impossible but possibly useful.

But outcome measurement in particular may carry another danger – the focus on outcome reduces the means to a merely secondary consideration. There are plenty of social problems that would be easy to ‘solve’ if there were no limitations on the means used to do so.

Puzzles and mysteries

Many of the issues that foundations address are more like mysteries than puzzles; in other words, there is no single solution and it is not known. This has a number of implications. One is that hard measures need to be kept in perspective. The second is that we need to leave room for serendipity.

As one respondent noted: ‘Before Christopher Columbus went to the New World he had no idea where he was going. When he got to the New World he had no idea of where he was. And when he came back from the New World he had no idea of where he’d been’.

Limitations and Future Research

There some constraints that limit the generalisability of the findings. The participants in the interviews volunteered to take part, and therefore demonstrated a level of engagement and interest in the study of philanthropy that inevitably influenced the discussions. The rich understanding gained from their deep and applied knowledge of the Australian philanthropic sector nevertheless remains only a snapshot of a point in time, and can provide no longitudinal view.

Future research in this area might develop a model or typology of approaches to performance measurement in foundations such that comparisons both over time and with other countries may be made. Addressing the difficult issue of terminology and working towards an agreed set of definitions would also assist in both the study and the practice of performance measurement.

And Finally...

This report offers an exploratory study of Australian foundations’ approaches to performance measurement for practitioners, policymakers and scholars. It draws on both academic and practitioner literature to provide a framework for the voices and insights of those most intimately connected with Australian grant-making foundations: their staff, managers and board members. In so doing, it reflects the importance of these foundations and their beneficiaries in our national economy and society.
References


Appendix One: Literature Review

The importance of audit, evaluation, and measurement of various types is taken for granted in broad swathes of economic and social life, transcending borders and contexts (Power, 1994, 2009). But are there some fields and institutions that are immune, or at least resistant, to measurement of performance?

One possible area of exemption would be the nonprofit sector. In economic and political theory (for example, Hansmann, 1987; Kingma, 1994; Weisbrod, 1987) the nonprofit sector’s very ‘nonprofit-ness’ is its badge of trust. Without profit, without anything to gain from under-performing or cutting corners, so the theory goes, nonprofits can be trusted to perform well and measurement is unnecessary. It might further be suggested that the nonprofit sector would be exempt from measurement because it is, in theory, based on altruism and compassion and we do not typically ‘assess’ virtues.

Philanthropic foundations’ approaches to measurement are particularly interesting and puzzling because, in addition to the issues noted above concerning nonprofit organisations in general, grant-making philanthropic foundations with a permanent endowment do not need to ‘perform’ in order to survive. The only ‘performance’ that materially affects their futures is the performance of their investment portfolio. But foundations are being urged to ‘measure’. As Katz noted in 2005: ‘The question is no longer whether to measure but how to do so constructively’ (Katz, 2005, p. 127, original emphasis).

Grant-making foundations typically depend on others, their grantees, to ‘do the work’ while the foundation is responsible for carefully selecting the grantees and then writing the cheques. So when a foundation measures, why is it doing so, who/what is it measuring and for what purposes? Measuring is additionally puzzling in foundations because a foundation typically pays the full amount of a grant upfront before the ‘performance’ even begins. And, even if it does not do so, an obligation to repay any portion of the grant not used or required for the project or purpose runs the risk of the grant being subject to GST. Given that most foundations dislike spending money on ‘administration’, why then would they invest in obtaining ‘out of date’/post facto information on a grant that (a) has already been paid and (b) has no material effect on the foundation’s organisational survival?

The issue of foundations assessing grantee performance seems to be more commonly discussed in the literature than foundations assessing their own performance. Both are considered separately below.
Foundations Assessing Grantee Performance

Several reasons why foundations measure grantees’ performance might be suggested including:

- To ensure proper use of money (accountability)
- To enable adjustments during the course of the grant
- To assist in (future) best use of money by grantees
- Learning – for future grant-making by the foundation
- Learning for replication/practice in the wider philanthropic sector (other foundations)
- Normative/professional expectations
- Reputational gain/marketing for the foundation and for the grantee
- Personal and organisational satisfaction that the ‘right’ choices have been made.

Foundations Assessing Foundation Performance

Why would a foundation measure its own performance? There are several possible reasons:

- Accountability - but this raises questions of to whom: donor, governors/board, regulator, shareholders, employees?
- To publicly demonstrate that money has been well spent. This is likely to be relevant to community and corporate foundations and to companies managing foundations that are wishing to secure or increase resources. It is not clear why fully endowed foundations need to do this, other than for self-satisfaction, and perhaps to demonstrate leadership in the field.
- Reputational gain/marketing (see note above).
- To assist in (future) best use of money/learning – for future grant-making
- To enable adjustments in foundation process and practices

Foundations are generally bigger than the sum of their grants, doing a variety of things outside traditional grant-making (such as providing technical assistance and expertise, publications, publicising research results, convening, lending profile and credibility to an issue etc.). It is difficult to develop similar measures to be used across all program areas, especially given broad missions and goals and lack of support and allocation of funds for organisational learning. Putnam (2004) provides some useful examples of techniques and foundations that are developing new methods, plus a short bibliography.

Ostrower (2006; see also 2004) also highlights the fact that ‘effectiveness’ takes different forms in foundations’ eyes. Components of effectiveness include: good grant-making/strong grant quality,
attaining foundation goals, financial performance, helping a geographically defined community, and having an impact (Ostrower, 2006). Presumably, different definitions of effectiveness would require different types of measure. But, this consideration may be some way off - as Ostrower notes: ‘The generality of the responses is indicative of the fact that so many foundations have not really thought through the specific meaning of effectiveness ...’ (Ostrower, 2006, p. 3).

Some foundations might argue that if they measure the performance of grantees then they have in effect measured their own performance. Aside from questions of cost and value-added by foundations the argument that foundation performance is the sum of the performances of their nonprofit grantees may be objected to on the grounds that ‘Foundations have a limited claim on the accomplishments and value of the nonprofit sector’ (Cohen, 2013).

Nevertheless, according to Boris and Winkler, there is a sense of momentum and commitment to charting impact among some of the largest US foundations even if the approaches they use are diverse and the language is idiosyncratic. However, they also note: ‘few foundations implemented internal processes to measure their own performance, or invested in evaluating their grant-making programs’ (Boris & Winkler, 2013, p. 69).

There are no comparable data from other parts of the world, but these findings from larger US foundations may serve as a useful background to exploration of the Australian scene.

Drivers of the Rise of Measuring Performance

Existing research (see, for example, Barraket, 2008; Considine, 2001; Hood, 1995; Leat, 2006; Nicholls, 2009) provides a range of explanations of the drive for measurement of nonprofits/grantees’ performance. Government, donors, trustees, regulators and taxpayers are all alleged to be making new demands of nonprofit organisations in general (see for example, Ebrahim & Rangan, 2010).

More specifically, sources of the drive for measurement of grantees’ performance might be attributed to: funders wanting to see that they are making a difference and making best use of their monies; committed nonprofit leaders and social entrepreneurs wanting better and more innovative solutions; growth and professionalisation of the sector; the growth of common administrative norms and use of common grant-making software; and the growth in credentialed experts in measurement.

There is much less discussion of pressures for foundations to measure their own performance and interestingly many of the explanations relating to nonprofits/grantees do not obviously apply to foundations (e.g. other than tax payer subsidy, foundations do not typically receive large sums of public money). There are occasional, anecdotal suggestions that new, younger donors with
backgrounds in the corporate sector are driving the push to measure, but arguably the majority of major foundation creators have always come from business (Carnegie, Rockefeller, Sage, and so on in the US; Potter, Myer and Fairfax in Australia) so some additional explanation is required.

Over the years the language has changed from research and evaluation in the 1980’s, planning and evaluation in the early 1990’s, and organisational learning in the late 1990’s, to ‘impact something or strategic something’ in the 2000’s (Patrizi quoted in Boris & Winkler, 2013). Interestingly, although these movements and approaches to philanthropy have been presented as something new, Katz (2005) argues that they are not so new at all - thus making explanations of the push for measurement in the foundation world even more of a puzzle.

Intended Purposes, Uses and Benefits

In the research literature the supposed purposes/benefits of measuring vary within the nonprofit sector. For example, ‘some nonprofit managers view performance measurement as a strategic management tool, others simply view it as a marketing and promotional tool, and still others regard it as a “resource drain and distraction”’ (Carman & Fredericks, 2008, p. 51). It is suggested that performance measurement can help nonprofits ensure survival and success within the sector; better manage resources and services; in their marketing efforts; in strengthening relationships with stakeholders, and in conveying their record to current and prospective clients, donors, board members and staff, and volunteers (LeRoux & Wright, 2010).

It may be important to distinguish between ‘intended purposes/benefits’ and ‘benefits in practice’. An organisation may undertake certain activities with the purpose of achieving x, but then either not achieve x, or achieve x but fail to build on/implement that. The obvious example in this context is that of a foundation asking for performance reports from its grantees, receiving them, and then never reading or using them.

Findings on lack of use are reported in a range of studies. Evaluations do not appear to be used for program management and strategic planning (Buteau & Buchanan, 2011). Braverman, Constantine and Slater (2004) suggest that most foundations do not regularly make use of evaluations. Kramer and Bickel (2004) note that ‘few evaluations appear to produce significant changes in the behaviour of foundations ... or grantees’. Patrizi and McMullan (1999) report that foundation boards are generally not satisfied with the evaluation products they receive due to absence of conclusive findings and clear evidence that outcomes can be attributed to their funding. Greenwald (2013, p. 5) suggests that ‘few individual grantees appear to be defunded as a result of evaluation’. Similarly, ‘in philanthropic
relationships allocating funding on the basis of measurable results and withdrawing support for poor performance is shaped by other considerations.’ (Benjamin, 2010, p. 393).

One explanation of this disjunction between practice and application may lie in Frumkin’s (2006) observation that donors are motivated by expressive motivations and not by evidence of what works. Some foundation decision-making may not accord with the picture of rational pursuit of outcomes.

If measurements are not typically acted upon why engage in the practice? DiMaggio (2001) suggests: ‘The good thing about the quasi-religious quality of rationality enhancing techniques such as performance assessment is that even when the techniques fail, the ritual of their application, like any successful ritual, may move people to self-improvement and cooperative action and so enhance the capacity of the organizations or industries to achieve collective goals’ (p. 250).

**Techniques and Sources**

Tools and Resources for Assessing Social Impact (TRASI - a tool launched in 2010 by the Foundation Center in the US) reveals 150 different approaches to measurement used by foundations, but where foundations obtain knowledge of methods and how they assess one as better than another is not clear (see also Maughan (2012) for an analysis of methods in use in Australia).

In recent years social return on investment (SROI) has been widely discussed. SROI attempts to measure and blend economic returns, socio-economic returns and social returns all expressed in monetary, quantitative or qualitative terms. However, SROI is still at a very early and imperfect stage of development and is criticized on a number of grounds including ‘adventurous’ calculations of ratios, lack of standardised data bases for financial proxies and failure adequately to capture the social element (Krlev, Münscher, Mülbert & Mildenberger, 2013). SROI is also subject to many of the criticisms levelled at goal oriented approaches to performance measurement discussed below.

**Costs and Criticisms of Different Approaches**

Criticism of existing performance measurement is not hard to find. Approaches are said to be cumbersome, complex, imperfect, and often flawed. In the nonprofit world they have traditionally focused on inputs and outputs, in part perhaps in an effort to deal with the fact that the goals of nonprofits are often vague, value-laden and diverse (not least because of the variety of stakeholders). In addition, there has been lack of knowledge about relationships between measures and goals (Buckmaster, 1999, p. 187; see also Ostrower, 2004). General criticisms also include the costs for the foundation and the grantee, the uncertain use to which reports are put, and the general lack of measurement capacity in the sector.
On closer examination, however, different approaches may have different strengths and weaknesses.

**Goal oriented approaches**

Evaluation on whether formally stated goals are achieved – takes for granted that it is meaningful to talk in terms of inputs, outputs and processes, and long term sustainable outcomes (or, in some definitions ‘impact’) are seen as the ultimate justification of any activity. Programs are nothing more than tools to outcomes.

But there are at least five criticisms of goal oriented approaches:

1. Causality is not considered, and so with no link between intervention and effect, goal oriented approaches attribute outcomes in a relatively random way.
2. Goal-oriented approaches have very limited formative use – they are not useful for making recommendations on how to improve performance.
3. Side-effects (either positive or negative) are not considered. Given that most social problems are multi-dimensional and interlinked and interventions tend to be organised as complex cooperative processes: ‘it is somewhat paradoxical that one of the dominant steering regimes of our time recommends an intensified exclusive focus on narrowly defined outcomes.’ (Dahler-Larsen, 2013, p. 15)
4. Monitoring of outcomes may change the activities being measured e.g. teaching to the test, labelling achievers, goal displacement etc. This may be true of other approaches but because of the focus on goal achievement as the key criterion this criticism is especially relevant.
5. Policy goals are often broad and ambiguous so, in order to measure goal achievement, standardised program outcome measures are needed. Quality may then be defined by ‘measurability’ rather than by the policy goal. The danger, of course, is that the tail of measurability begins to wag the dog of policy goals.

**Theory based approaches**

Theory based approaches apply an underlying general set of assumptions, often called program theory to explain why activities can plausibly lead to a specified outcome and what else is likely to happen.

The advantages over goal based approaches include a focus on how and why the program works, pinpointing links in the causal chain and where repairs may be needed. Theory-based approaches can also distinguish situations where the problem is poor implementation from those where the underlying theory is flawed and it may be better to terminate the program. In other words, theory
based approaches can distinguish implementation failure from theory failure – the former may require more resources, the latter may require termination.

Theory based approaches also open a way to quality outcome measures, avoid holding people to account for outcomes not under their control, allow for recognition of side-effects, and allow knowledge to be accumulated over several evaluations. But again, there are criticisms. Theory based approaches are said to:

1. Require large amounts of time
2. Require not one program theory but several depending on the range of stakeholders and assumptions
3. Be flawed in that any given practice is more complex than a theory of it – practice is sensitive to specificities of time, situation, person etc. so an explicit program theory is a reduction of practice not a representation of it.

Participatory and responsive approaches

Participatory and responsive approaches invite various participants to share responsibility for the evaluation process to enable construction of contextually meaningful knowledge. One advantage is said to be that this promotes ownership of evaluation results, and learning. But inevitably, these approaches have their critics too:

1. If utilization has top priority then evaluation will be primarily concerned with small adjustments.
2. Which stakeholders and stakeholder’s views have priority?
3. How to deal with changing and evolving goals?

For further detail on the limitations of different approaches see, for example, Dahler-Larsen (2013), and Forti and Yazbak (2012).

Conclusion

The very limited operational scope of much nonprofit work has implications for outcome measurement. ‘Measuring impacts makes sense under a severely limited set of circumstances – when an organisation operates at an ecosystem level, and yet can exercise sufficient control over results to attribute impacts to its work.’ (Ebrahim & Rangan 2010). This is not the typical scale of the vast majority of nonprofit organisations and foundation grants. Ebrahim and Rangan (2010) also note the tension for nonprofits between a focus on measuring only what is easily measurable (inputs and outputs) and overreaching by taking credit for social changes beyond their actual control (p. 13).
Benjamin (2010) highlights the ways in which philanthropic relationships are not equivalent to accountability relationships where a principal hires an agent with desired competencies and then rewards that agent for results (known as agency theory). In philanthropic relationships other considerations including the capacity of a grantee, limited resources of a grantee and grantee developmental or relationship goals may be at odds with short term results. ‘The philanthropic relationship may be better characterized as a relationship of mutual interest that addresses a common concern rather than a relationship of delegated authority’ (p. 392).

Because of these differences funders that adopt performance measurement may encounter tensions as these new practices challenge and/or crowd out the importance of other philanthropic concerns (such as supporting grantees, capacity building etc.). Without a way to address these concerns, Benjamin argues, funders face a choice between prioritising measurable results and falling back on funding good intentions.

In philanthropy, as in agency theory, there are potential problems of moral hazard and information asymmetry, or imbalances of risk and information between the funder and the grantee. Although foundations do not want to fund incompetence they may want to support nonprofits with limited capacity. So, Benjamin concludes, we need to understand how the characteristics of the philanthropic relationship – longevity, depth, breadth – shape a funder’s use of performance measurement and if and how capacity concerns are dealt with.

So in putting together the academic research literature with the far more extensive professional literature on ‘how to measure’, there is a disjunction. Advice on how to measure abounds, discussion of why and for what purposes and benefits is scarce. What does seem clear is that evidence from performance measurement is not the only factor in decisions made by foundations and by other donors.
References for Literature Review


### Appendix Two: Meanings and Standards of Evidence

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impact</strong></td>
<td>The set of practices through which an organisation establishes what difference its work makes. ...includes measuring both outputs and outcomes. This does not mean we equate measuring outputs with measuring impact—it is just part of the overall process.</td>
<td>Assesses what has happened as a result of the intervention and what may have happened without it, from a future point in time.</td>
<td>The long term outcomes that may be attributed to the grant (and any associated activities) adjusted for what would have happened anyway, the actions of others, drop-off and displacement.</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td>The results of a programme or intervention.</td>
<td>Those benefits of changes for individuals or communities after participating in the programs of nonprofit organisations.</td>
<td>An assessment of the results of a program activity compared to its intended purpose.</td>
<td>The changes, benefits, learnings or other effects (both long and short term) that result from the organisation’s activities.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outcomes may be initial (immediately after); intermediate (after 12 months); long-term (after 12 months and further over extended periods).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Outputs</td>
<td>Direct products of program activities.</td>
<td></td>
<td></td>
<td>The tangible products and services that result from the organisation’s activities.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>The use of information from monitoring and elsewhere to judge and understand the performance of an organisation or project.</td>
<td>Objective assessment of an ongoing or recently completed project, program or policy, its design, implementation and results.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monitoring</td>
<td>A systematic way of collecting and recording information to check progress against plans and enable evaluation.</td>
<td>Regular systematic collection and analysis of information to track the progress of program implementation against pre-set targets and objectives.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are numerous other definitions; for example, Fiennes runs together evaluation and impact assessment suggesting that evaluation ‘involves investigating whether a decent breakfast actually does aid learning. And that requires isolating the effects of the intervention (providing the breakfast) from other extraneous factors. This is impact evaluation’ (Fiennes, 2013).

Some organisations in the US and elsewhere set very high standards of evidence. In addition to different definitions of outcomes, evaluation, impact and so on there are different standards of ‘evidence’ of effective performance. For example, Intelligent Philanthropy suggests that five questions constitute evidence of a charity’s effectiveness. Does a charity:
1. Track outcomes years after serving people, or only annual accomplishments?

2. Measure program outcomes against relevant benchmarks, averages or control groups?

3. Complete independent evaluations of program effectiveness?

4. Survey beneficiaries about program quality and impact?

5. Pursue specific and measurable organizational goals?

If a charity is not taking all five steps to track outcomes properly, then you don’t know if their work is effective or not. I fear the majority of charities could not answer all five questions affirmatively..... That does not mean those charities are failing to create lasting improvements, but it does mean we will never know one way or the other’ (Penley, 2013).

The Edna McConnell Clark Foundation adopts a different position distinguishing between: (i) ‘proven effectiveness: experimental research has confirmed the program’s impact on participants; (ii) demonstrated effectiveness: systematically collected data comparing program participants with similar people not receiving a program’s service enable an organization to substantially conclude that youth are benefiting from program; (iii) high apparent effectiveness: systematically collected data indicates youth are probably benefiting as intended from participating in a specific program’.

(http://www.emcf.org/our-strategy/our-selection-process/evidence/)
Appendix Three: Interview Questions

Study of Approaches to Performance Measurement Among Australian Foundations

Foundation Interview Schedule

A few points before we start. First, in this study we are more interested in your broad perspective on performance measurement than in the nuts and bolts of what you do or the results you achieve. Second, we are using the term ‘performance measurement’ very broadly to encompass any form of evaluation, analysis, data collection etc. designed to assess results. Third, we want you to focus in your answers on your current practice (rather than what you used to do).

Section 1

Background

1. Name of foundation and your role
2. Age
3. Annual grant-making spend
4. Total annual spend

(all of above from Annual Report/Directory of Phil; then check with interviewee)

I would like to start by hearing a little more about your foundation and its work:

1. Mission
2. Style: grant-making, operating/funder plus, mixed
3. Key programmes
4. Pattern of grant-making – % grants
   - $1,000
   - $1,001 - $10,000
   - Between $10,000 and $50,000
   - Over $50,000
Section 2

Drivers of the rise of performance measurement

1. Do you feel that Australian foundations are under pressure to evaluate or measure the results of their grants?

2. Do you feel any pressure in your foundation to evaluate or measure the results of your grants?

If NO: go to q 17 and on

If YES:

3. How strong do you feel that pressure is?

4. When did you first notice the pressure – less than 5 years ago, 5-10 years ago, more?

5. Do you see the pressure as getting stronger or fading?

6. What do you see as the sources of the pressure?

Prompt: foundation peers, trustees/Board, government/regulators,

Other: declining income/doing more with less, frustration at lack of change, frustration at charity results, anxiety about foundation reputation and/or legitimacy.

7. Does the pressure apply across the whole foundation or do you see it as only applying to specific programmes and/or grants?

8. If only to specific programmes/grants – which ones and why the variation?

9. Roughly what proportion of your grants by number is evaluated, measured etc.?

   Less than 20%, 20-50%, 50-80%, 80-100%

10. AND – do you also measure the foundation’s own performance?

If YES: how do you do this?

If NO: why not? What are the obstacles?
Section 3

*Purposes of performance measurement*

11. When you talk about measurement what language do you most often use?

Prompt: evaluation, performance measurement, impact measurement, SROI,...

And what does it involve? (an acquittal report, an independent evaluation or ??)

12. What does your foundation see as the main purpose(s) of measurement?

Prompt: ensuring the money was properly spent, learning, knowing what works, knowing what doesn’t work, informing future grant decisions, accountability...

13. Do you think everyone in the foundation sees the main purpose(s) in much the same way?

Prompt – board, staff, CEO

14. and what about grantees – how do you think they see the main purposes?

15. If the performance were disappointing what would follow?

16. If the performance were excellent what would follow?

Probe: so have you ever decided to 1. Continue or 2. Adjust or 3. Terminate - a grant on the basis of an evaluation?

Section 4

*Focus, design, responsibility*

17. When is the performance measurement planned – before or after the grant award?

18. Who is responsible for:

   Making decisions about whether to measure?

   Study design, metrics and timing?

   By whom the work (measuring) will be done?

   Documenting and disseminating the results?

19. Is there a particular person in the foundation responsible for overseeing performance
measurement?
If YES: expand

20. Do you do any collaborative funding?
If YES: what effect does this have on your approach to measurement?
Prompt: is it negotiated? Who leads? What are the criteria (e.g. biggest $ contribution?)

Section 5

Techniques and their sources

21. How has your foundation learnt about performance measurement approaches?

22. How do you assess one approach as more suitable/fit for purpose than another?
Prompt: size, proportionality, geographical location

23. Are there any methods you find particularly useful?
Prompt: figures, cases, histories...

24. After how long would you typically make a final assessment of performance?

25. How do you allow for the effects of external factors on results?

26. How do you allow for what might have happened anyway i.e. even if you had not made the grant?

++ probe for attribution/contribution

Section 6

Costs and benefits

27. What do you see as the direct and indirect costs and benefits of measurement for the foundation?

28. What do you see as the direct and indirect costs and benefits of measurement for the grantee?

29. What do you see as the direct and indirect costs and benefits of measurement for others?
Section 7

Effects of Performance Measurement

30. Do you feel that measurement/evaluation etc. have affected your appetite for risk?

31. Has measuring the performance of or in programmes had any effects on priorities/the programmes you work on?

Section 8

More or Less

32. In a year’s time do you think you will be doing more or less measurement?

If MORE: why do you say that? If LESS: why do you say that?
Appendix Four: Research Team

Professor Diana Leat – Ian Potter Foundation Visiting Scholar 2013, ACPNS

Diana is currently a Visiting Professor at Cass Business School, London and at the Australian Centre for Philanthropy and Nonprofit Studies, QUT Brisbane. Diana has worked with the Carnegie Trust in the UK developing a consortium of social justice oriented foundations, the first research centre for philanthropy in the UK (CGAP), and compiling a series of case studies on social change philanthropy. She is the author of over 100 articles and books on the nonprofit sector and social policy, specialising in philanthropic foundations. Diana was a trustee of the Diana, Princess of Wales Memorial Fund in the UK, and a member of the NCVO Advisory Council. Diana’s QUT research publications can be found at http://eprints.qut.edu.au/view/person/Leat,_Diana.html

Alexandra Williamson

Alex is a part-time Research Assistant in the philanthropy team at the ACPNS, and is currently also studying for a Masters by Research, looking at the accountability of endowed Australian philanthropic foundations. Prior to joining QUT she worked in the research grants office of a Go8 University, and in several grant-making roles in private foundations and trustee companies. She holds a Master of Business (Philanthropy and Social Investment) from Swinburne University of Technology. Alex’s research publications can be found at http://eprints.qut.edu.au/view/person/Williamson,_Alexandra.html

Dr Wendy Scaife

Wendy is Senior Research Fellow and Acting Director of the Australian Centre for Philanthropy and Nonprofit Studies at QUT. Wendy studies and teaches philanthropy and fundraising and has a 25 year involvement in the philanthropy and nonprofit sectors. Her 12 years of research work spans the cause/giver spectrum. Wendy was on the Research Australia Philanthropy and National Compact Steering Committees serves on FIA’s Codes Taskforce and Professional Development Committee, and on the research committees of the Charities Aid Foundation and the US Association of Fundraising Professionals. She is on a grants committee for the Foundation for Rural and Regional Renewal and also chairs a community heritage project. Wendy was a state CEO/ national Deputy CEO of the Leukaemia Foundation. Wendy’s research publications can be found at http://eprints.qut.edu.au/view/person/Scaife,_Wendy.html