3 A measurement framework

Key points

- Interest in measuring the contribution of the not-for-profit (NFP) sector has been prompted by:
  - the sector's increased share in economic activity, increasing appreciation of the social benefits it delivers, and its greater engagement by government in service delivery
  - demand by NFPs to measure their contribution in order to improve organisational performance and demonstrate the merit of their activities to philanthropic and government funders.
- This chapter develops a framework, based on impact mapping, to guide and improve measures of the contribution of NFPs. There are four distinct levels at which contribution can be measured:
  - *inputs* (measures of the resources used)
  - *outputs* (indicators of the level of activity undertaken)
  - *outcomes* (direct costs and benefits to the activity participants)
  - *impacts* (longer-term net benefits to the participants, and other costs and benefits to the broader community).
- Contribution can be measured for the entire sector, for groups of NFPs pursuing similar goals or for individual organisations, programs or activities. Input and output measures provide the only consistent basis on which the contribution of the entire sector can be measured. Measuring outcomes and impacts is particularly important for understanding organisational, program and activity effectiveness.
- NFPs contribute through:
  - *service delivery* to members or clients
  - *exerting influence* and initiating change in economic, social, cultural and environmental issues
  - *connecting community* and expanding people's social networks
  - *enhancing community endowment* by investing in skills, knowledge and physical, social, cultural and environmental assets for current and future generations.

  NFPs may pursue one, some or all of these purposes and their outcomes can interact with others in shaping the eventual impact.
- The proposed framework can accommodate the full range of measurement techniques applied to estimating the outcomes of NFP activities. The merit of each technique will depend on the purpose for which measurement is undertaken and the information available. As such, there is no 'gold standard' approach to measurement which will be best suited to all NFPs under every circumstance.

3.1 Why measure the contribution of the sector?

There is growing interest in the sector

Interest in measuring the contribution of not-for-profit organisations (NFPs) has grown over the past decade. An internationally-accepted statistical classification for the sector has been developed (Salamon et al. 2007) and a range of other measurement and evaluation tools has been designed, or adapted, for use by NFPs.

As noted also by VCOSS (sub. DR276), this interest has been prompted by:

- greater recognition of the sector's importance in community life and economic activities (UN 2003)
- attention from researchers and policymakers on the sector's role in achieving social and economic objectives (for example, its role in promoting community cohesion, advancing democratic principles, and providing the opportunity for spiritual expression)
- increased recognition by NFPs of the benefits of evaluation and measurement in improving their organisational performance
- demands by donors and governments for improved performance evaluation to enhance accountability and provide more information on the effectiveness of their funding
- governments seeking greater understanding of the structure of the sector in order to assess the potential impact of changes in policy such as regulatory arrangements, tax concessions, or programs to support volunteering.

Reflecting these various motivations, measurement has been undertaken from two perspectives — the 'macro' level (covering the entire sector or significant groups within it) and the 'micro' level (focusing on the organisation, program or activity).

This chapter outlines a framework for measuring the contribution of the NFP sector that can be used to guide reporting at both the 'macro' and 'micro' levels. The proposed framework is broad enough to capture the many facets of NFPs and flexible enough to enable users to choose measurement techniques which best suit their own circumstances. Participants to this study have identified a range of benefits from measurement (box 3.1). They point to the need for a measurement framework that supports the evaluation, as well as reporting, of contribution.

Box 3.1 The value of measurement: views of study participants

Measurement can improve organisational performance

Mission Australia has made, and is making a significant investment in quality assurance, systems support and other measurement and evaluation tools supported by research, to track and ensure its programs are effective in delivering outcomes for its vulnerable and disadvantaged client group. (Mission Australia sub. 56, p. 1)

Nonprofit organisations are … seeking to improve their management and governance and to reposition themselves in an increasingly competitive environment. Their leaders are looking for data on the average performance of nonprofits in similar industries to provide them with benchmarks against which to assess their own performance. (Lyons sub. 169, p. 8)

Measurement provides information to funders

In order to fulfil their missions and deliver on impacts, nonprofit organisations depend on revenue from grant making bodies, public fundraising and, where applicable, membership bases. Reporting transparently on outcomes and impacts is essential to safeguarding the trust of these stakeholders. Developing a system to effectively measure the contribution of organisations and the sector as a whole would facilitate and strengthen the relationship between organisations and their support bases. (Fundraising Institute of Australia sub. 76, p. 4)

Measurement provides insights into the operation of the sector

… attention needs to be place on the measurement of the broader contribution of the sector. Of particular importance is a greater awareness of the contribution of the formation of networks, sometimes between individuals and sometimes between organizations. We know relatively little about how these operate in providing broad communication channels and support within civil society, initiating new social movements, or raising issues of concern in the community. (Professor Jenny Onyx and Jenny Green sub. 13, p. 1)

It is crucial in discussions around measurement and performance accountability of community based organisations to consider the evidence of the distinctiveness of their practice. It is not possible to measure contribution without an understanding of the nature of the contribution and how it is achieved. (Illawarra Forum Inc. sub. 52, pp. 5–6)

Measurement can improve public policy outcomes

… a nationally consistent framework for the measurement of the sector’s contribution that includes **agreed performance indicators to standardise reporting arrangements** would aid comparative evaluation of the efficiency and effectiveness of various services. This would result in more optimum decision making by governments in respect of resource allocation and help to deliver the best possible outcomes for the community. (Emphasis in original) (Communities@Work sub. 150, p. 8)

Governments have long relied on nonprofit institutions for the provision of many services. Understandably, their focus has been exclusively on those organisations they fund ... Reflecting overseas trends, Australian governments are coming slowly to recognise the wider contribution of all nonprofit organisations, especially their overall contribution to sustaining strong communities. They have sought to encourage volunteering and philanthropic giving. Such policy initiatives need reliable and regular data if their effectiveness is to be evaluated. (Emphasis in original) (Lyons sub. 169, p. 8)

Contribution is measured by net benefit

Measuring contribution involves identifying all community impacts from a particular activity. This differs from measuring the level of economic activity, which is a partial measure of contribution as it only captures market based activity, valued at market prices. To fully measure contribution, measures should include all economic, social, cultural and environmental costs and benefits accruing at the individual, group or broader community level. These should include the costs and benefits associated with broader, including unintended, consequences, as well as for those directly involved in the activity.

Net community benefit is measured relative to the situation which would have existed had the activity not taken place (that is, against the 'counterfactual'). It is important to emphasise that not all sector activities necessarily deliver a net community benefit. This point is particularly relevant in circumstances where the interests of members diverge from those of the broader community. According to Lyons:

> There is, however, a dark side to nonprofits. Sometimes nonprofit organisations are formed to separate their members from the wider society, sometimes to denounce or object to other groups. Democratic politics requires a clash of organized interests, but it also requires that those interests exercise a degree of self-restraint. Some nonprofit associations ignore this and cause wide social division and conflict. In some fields, nonprofits have been criticized as encouraging elitism or as embodying an outmoded charity that stigmatizes and degrades its object. Yet no considered assessment could view these flaws as any more than a small part of the total picture. (1998, p. 16)

Characteristics of a good measurement framework

Measurement serves a number of purposes. It provides information on resources used by the sector, the benefits (and beneficiaries) of NFP activities, the clients and members who are targeted by their activities and organisational structures within the sector. This information can be used to gauge how efficiently NFPs meet their objectives, the nature of their interaction with others in the community (including any additional or 'spillover' effects these may generate) and, for governments, provide an indication of the possible value in implementing public policy. It is also important for assessing the consequences of changes to the environment in which NFPs operate (for example, labour market conditions or the public policy environment).

Measuring the contribution of the sector gives rise to a number of challenges. These include the expense of undertaking measurement, the difficulty of measuring intangible contributions and producing comparable results, and the possibility that measurement may encourage organisations to focus on activities which are easier to measure rather than those which deliver the greatest social benefit (chapter 5).

These challenges emphasise the importance of developing a commonly agreed framework which:

- enables the use of statistics from different sources to provide a more comprehensive approach to measurement
- allows gaps, inconsistencies and duplication to be more readily identified and provides a basis for standardisation of some measures (such as using a Standard Chart of Accounts for defining input measures)
- can be used, or adapted for use, by any type of organisation within the sector (including those that lack the resources to undertake comprehensive measurement or evaluation)
- is sufficiently rigorous to enable meaningful comparisons to be made between organisations within the sector (and with organisations in other sectors)
- recognises that not all contributions may be able to be quantified
- is flexible enough to acknowledge the diverse activities, organisational forms and the (sometimes unique) contributions made by NFPs.

Impact mapping provides an ideal measurement framework

Impact mapping provides a logical structure through which the operational decisions of NFPs can be followed through to their ultimate impacts, either by the organisations themselves seeking to evaluate performance or by governments (or other funders) assessing the effectiveness of these organisations in program delivery. By explicitly identifying the links between the resources used by the sector, the activities undertaken and the subsequent results of those activities, the process of measuring can itself improve understanding of how the sector operates. And, by identifying the processes by which NFPs seek to deliver benefits, impact mapping also allows contributions to be identified in a qualitative sense where explicit quantification is not possible.

Input and output measures generally provide insights into the activities and processes of NFPs and, when aggregated, the scale and scope of the sector. Outcomes measures provide information about those who directly benefit from those activities (clients and members) while impact measures aim to reflect the net benefit for the broader community. In the case of NFPs providing international aid, many of the outcomes and impacts are in the development partner country. Given the diverse range of activities undertaken by the sector, and the consequent difficulty in producing comparable impact and outcome measures for aggregation purposes, these higher level measures are generally more suited to understanding the contributions of individual organisations or activities than the sector as a whole.

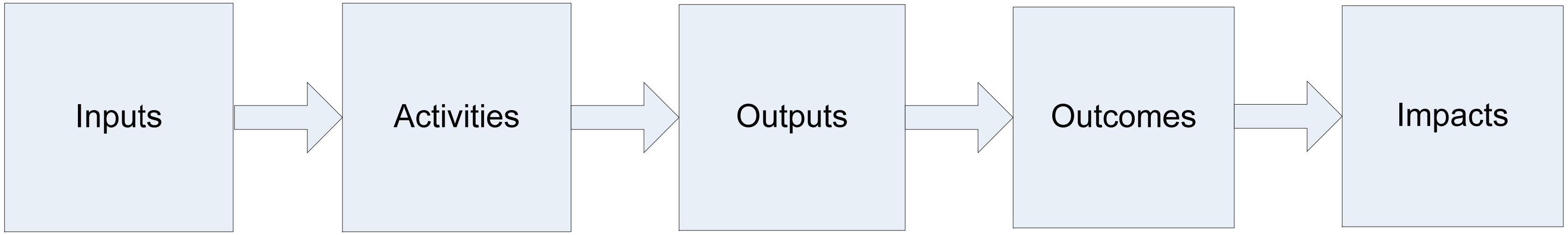
Impact mapping forms the basis of a number of evaluation methods — the OECD (2007) refers to this as a 'results chain'; the Social Return on Investment (SROI) refers to this as impact mapping, and ARACY refers to it as program logic (sub. DR199) (chapter 5). The core concept is mapping between the four levels at which contribution can be measures — inputs, outputs, outcomes and impacts. Figure 3.1 illustrates the levels of contribution. Input, output and outcome measures are commonly employed in program evaluation (United Way of America 1996; Atkinson Review 2005; OECD 2009b; SCRGSP 2008). The Atkinson Review uses the delivery of health services to illustrate how they differ:

> ... we identify *inputs* as the time of medical and non-medical staff, the drugs, electricity and other inputs purchased, and the capital services from the equipment and buildings used. These resources are used in primary care and hospital *activities*, such as a GP making an examination or the carrying out of a heart operation. These activities are designed to benefit the individual patient. To the extent that they do, the health care provided constitutes the *output* associated with these input activities. Finally there is the health *outcome*, which may depend on a number of factors apart from the output of health care, such as whether or not the person gives up smoking. [Emphasis in original] (Atkinson Review 2005, p. 40)

The final level — impacts — captures broader, longer-term effects (including those on the community). It therefore encompasses 'spillover' effects. To continue the analogy of medical services, impacts or spillover benefits would include the general productivity and social benefits of a healthier population.

Spillover effects, which may not have been intended by those conducting the activity, are sometimes incorporated into outcome measures. However, distinguishing between outcomes and impacts allows NFPs to evaluate the direct results of their activities (potentially enabling improvements to service delivery) separately from the longer-term effects and the broader consequences for the community. Measurement of change at the impact level is important to guide the allocation of resources to deliver the highest community wellbeing for the resources used. Disaggregated, 'impact' measures also provide information for targeting interventions as well as assessing broader trends in wellbeing.

Figure 3.1 **Impact mapping – levels of contribution**



Inputs

Resources used by the organisation

eg.

- Money
- Staff & staff time
- Volunteers & volunteer time
- Facilities
- Equipment and supplies

Activities

What the organisation does to fulfill its mission

eg.

- Feed and shelter the homeless
- Provide job training
- Educate the public about the signs of child abuse
- Counsel pregnant women
- Create mentoring relationships for youth
- Conserve historic heritage places
- Organise community events

Outputs

The direct products of the organisation's activities

eg.

- Number of classes taught
- Number of counselling sessions conducted
- Number of educational materials distributed
- Number of hours of service delivered
- Numbers of participants served

Outcomes

Intended benefits for participants during and after organisation's activities

eg.

- New knowledge
- Increased skills
- Changed attitude or values →
- Modified behaviour →
- Improved condition
- Altered status

Impacts

Longer term (including unintended) benefits for the individual and community

eg.

- Enhanced sense of self →
- Improved community participation
- Improved productivity
- Stronger communities

Constraints on the use of resources and delivery of activities

eg.

- Laws
- Regulations
- Funders' requirements
- Contractual service delivery requirements

Influenced by mediating environment

eg.

- Personal traits
- Economic conditions
- Social & demographic trends

Source: Based on United Way of America (1996). The figure has been adapted to include a separate 'impacts' category.

As indicators progress through the hierarchy from inputs to impacts they generally impart more comprehensive information for the purpose of assessing contribution. But, moving through the range of indicators in this way is also informationally demanding and less data is available to construct attributable outcome or impact measures than are available to construct input and output-based measures (chapter 4). In addition, outcomes and particularly impacts are usually the result of a much wider array of influences than just the NFP activity.

Sectoral level measures are generally constrained to lower level measures

Most measurement frameworks that seek to report at a sectoral level operate at only input or output level of contribution. The Australian Bureau of Statistics (ABS) *Non-Profit Institutions Satellite Account* — based on the *UN Handbook on Nonprofit Institutions in the System of National Accounts* (UN 2003) — provides input and output data for the sector. While only compiled for economically significant NFPs, the Satellite Account forms the first step of impact mapping — measuring the inputs of the sector and, in some cases, providing output estimates. Where output measures are not directly observable (such as through market transactions), estimates are based on the value of inputs used. This approach is utilised more generally in the National Accounts where output is difficult to measure, such as with government services (chapter 4).

While inputs, and outputs with market prices, are relatively easy to measure in money terms, they will understate the contribution of the NFP sector to wellbeing relative to the business sector if NFPs have relatively greater broader community benefits (spillovers). This makes measurement of outcomes and impacts important for informing resource allocation (chapter 2).

Case studies provide insights into the contribution of NFP's activities to outcomes and impacts. Meta analysis of a range of such studies improves confidence in the conclusions drawn and can provide benchmark measures that are representative of the sector more generally. As noted above, many of the sector's contributions are intangible in nature and hence not readily amenable to quantification, so only a subset of outcomes and impacts may be able to be 'valued' in dollar terms. As a result, a range of qualitative and proxy measures must be accommodated in the framework.

3.2 How do not-for-profits contribute to wellbeing?

NFPs contribute to community wellbeing in four broad ways:

- *Service delivery* either to people outside the organisation (such as social support or emergency services) or to members (this may include the opportunity to participate in worship through a religious organisation or the benefits of participating in a community club).
- *Exerting influence* and promoting change on economic, social, cultural and environmental issues. For example, NFPs influence public policy outcomes through advocacy on behalf of individuals (for example, clients of a disability support group) or on behalf of a community of interest (for example, a group with a shared interest in a recreational activity). Influence can also be exerted through education programs to change community perceptions, research to enhance understanding of issues impacting on the community, and demonstration of more effective approaches that then change approaches and resource allocation.
- *Connecting the community* and expanding the social networks available to individuals. Connection through worship, social and sporting clubs and other organisations that promote community engagement are obvious examples, but connection can also occur through volunteering, such as with a service delivery organisation. Provision of social support services can also provide a means by which socially excluded individuals can re-engage with society.
- *Enhancing the community endowment* by investing in skills, knowledge and physical, social, cultural and environmental assets for the benefit of future generations (for example: philanthropy which establishes art galleries or museums; environmental conservation; historical societies which preserve local traditions; or local community groups which maintain social networks that can be drawn on in times of crisis).

As 'mission' classifications these differ somewhat from some previous attempts to categorise the roles of the sector (box 3.2). The missions of some NFPs may result in activities which span all categories. Others may be more narrowly focussed on only one or two roles.

Box 3.2 Alternative mission classifications

Alternative mission classifications for the NFP sector have been proposed. Wolpert (2001) provided three classifications — 'philanthropy' (to establish and enhance civic institutions such as hospitals and social capital); 'service provision' (activities which foster mutual benefit, serve markets not served by the public or private sectors, and to enhance quality, variety, compassion and efficiency in service delivery) and 'charity' (redistributing resources to the needy). Land (2001) added an additional category 'fellowship' (to provide affiliation and association with other members who share an activity or interest).

The Commission's proposed classification encompasses all of these, but also explicitly recognises some additional primary objectives (such as the sector's educative role). It also avoids the potential confusion of identifying an input, philanthropy, as an ultimate goal in itself (rather than as a means to achieving other goals).

The classification categories allow for both direct effects (such as through the delivery of services to clients and advocacy for change policy) and indirect influences (for example, through building the capacity of participants to engage in other areas of economic and community life). They also support identifying primary outputs and outcomes and those that are secondary, (in some cases a by-product of the activity — such as the connections made from involvement in worship activities or friendships formed when volunteering for meals on wheels).

3.3 The proposed measurement framework

The proposed framework is outlined in figure 3.2. It is primarily a reporting framework to encourage the development of common measures and indicators at each level for each of the four categories described above. It also provides a structure to guide the user to articulate the relationships (draw the connections between) each level from the initial input decisions through to the ultimate outcomes for the individuals involved and the impacts on them and the community. Measurement in a consistent framework will promote data collection to support empirical testing of the theoretical links between levels or leaps of faith that underpin the development of programs and policies.

Different data collection and measurement challenges arise at each level in the framework. Methodological issues surrounding measurement of inputs, outputs, outcomes and impacts are discussed in more detail in appendix B.

Inputs

Inputs are the resources used by the sector. These include labour (both paid and volunteer), physical assets (buildings, equipment and food banks), and financial capital (which may be obtained from a number of sources including from government, philanthropists, member fees and investment income). Intellectual capital and the value of relationships (both within and outside the sector) are also important inputs, but can be particularly difficult to measure.

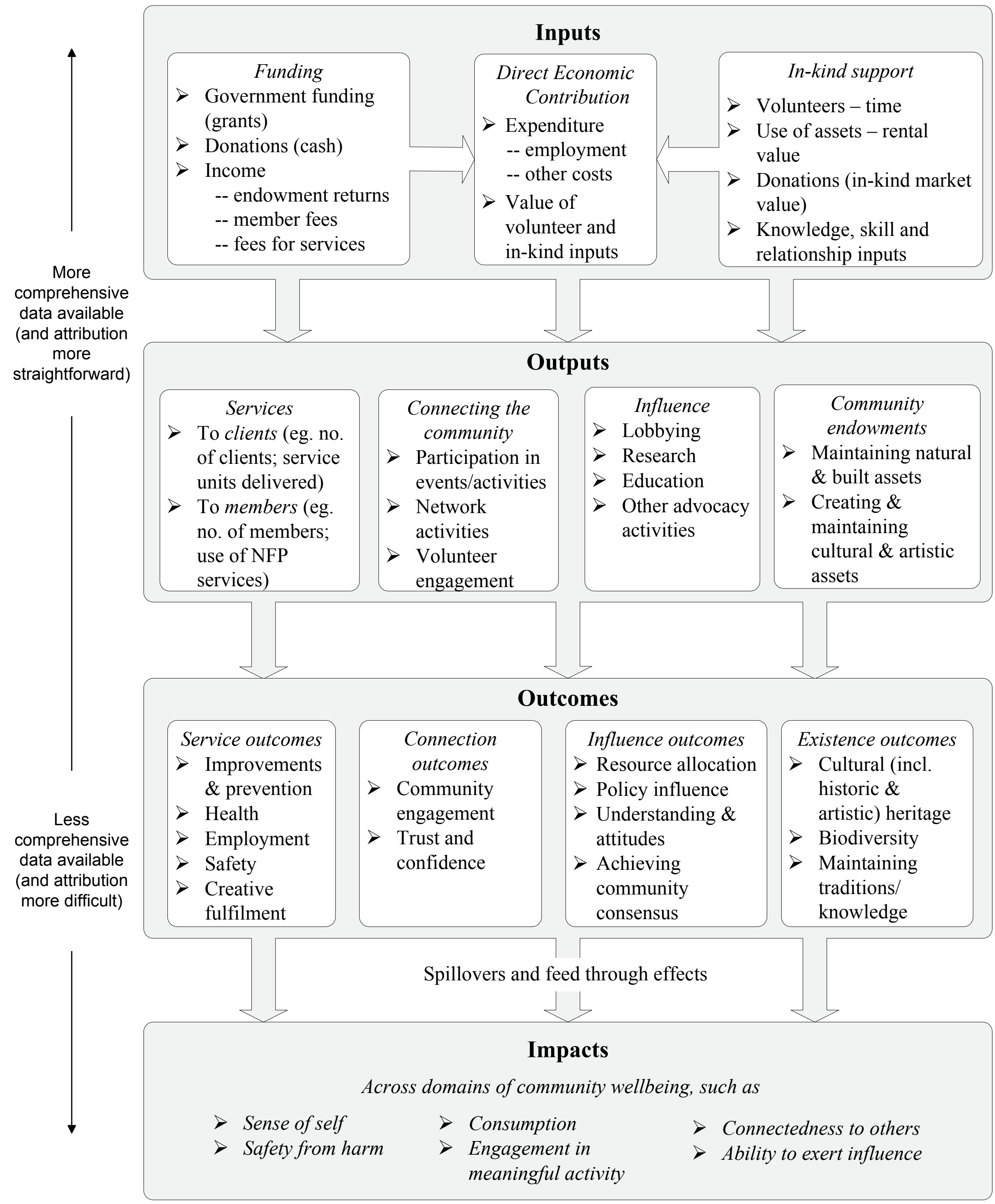
Input measures provide information on the processes by which activities are delivered and the types of resources used. They are also relatively easy to categorise by type of organisation and activity. Since contribution is measured by the net impact of an activity — the value of all benefits, including those to the broader community, less the value of resources used — input based measures are a necessary first step in evaluation.

Aggregated input data, for some parts of the sector, are available from a number of sources, including the ABS and the Australian Institute of Health and Welfare (AIHW) (chapter 4). Significant progress has been made in valuation of volunteer time to better reflect the contribution of skills. Adoption of Standard Chart of Accounts as proposed by COAG would substantially improve comparability of reported income and expenditure data (chapter 6).

Measurement challenges include the identification and valuation of in-kind contributions, and the estimation of capital assets and valuation of capital services. The measurement of intangible inputs such as trust, relationships and knowledge of the client population are particularly challenging, but here too progress is being made.

Input based measures provide the most reliable option for establishing a time series picture of the evolution of the overall level of activity and structure of the sector. If they are compiled according to an international standard, they also allow comparisons across countries. Moreover, although an incomplete measure, in some circumstances input measures can provide a reasonable proxy measure of outputs. The labour intensive nature of much of the sector, and the absence of a need to pay a profit, can make the cost of labour a reasonable approximation to the 'price' of output (Salamon and Dewees 2001).

Figure 3.2 **Measurement framework and types of indicators**



Outputs

Outputs are the direct result of NFP activities and processes. These outputs are not ends in themselves, rather they are a means of delivering outcomes. Output measurement should support tracking the connections between inputs and outcomes, as well as accountability. Both primary and secondary outputs should be identified, including incidental outputs where they are significant.

Consistent with the classifications outlined above, the framework categorises the outputs of NFPs into:

- services delivered to clients or members — for example, measures of services delivered include the number of clients treated for drug dependency, beds provided in emergency accommodation, and concerts performed by a community orchestra
- engagements which facilitate the connection of community members — for example, measures include community participation in sporting matches and the number and extent of volunteer involvement
- influence services, including education and advocacy — examples of measures include research papers, media coverage, training programs and lobbying materials
- investments in protecting, maintaining or building the endowment of social, cultural and environmental assets — for example, measures include historic heritage places conserved, and kilometres of riverbank fenced.

Outputs can be measured at the organisation (for example, number of clients served) or program level (for example, number of program participants). In addition to data contained in the ABS satellite account on the economic contribution of NFPs (chapter 4), output measures may be obtained from annual reports of NFPs and from government program reviews.

Comparisons across different types of output measures require volume measures (measuring number of services) to be converted into value, or dollar, measures. Some sector outputs have a market price (such as fees paid by clients), but most can only be measured by indicators such as membership or numbers participating in activities. The 'value' placed on these outputs will clearly differ by the type of organisation, activity and quality, and finding appropriate 'market prices' is a challenge in constructing the sector-wide output measures contained in the satellite account for the sector (chapter 4).

The Atkinson Review addressed this issue in relation to measuring the output of government, which faces similar measurement issues. Comparability requires that private and public sector outputs be measured on the same basis:

> The thrust of the [system of national accounts] was ... 'to treat, as far as possible, public output in the same way as private output: the same general procedure can be used in both the public and private sector.'... This seems clearly right. The issues of measuring output and productivity apply across the national accounts as a whole, and the

> principles applied to their measurement should, as far as feasible, be the same. This is particularly important in view of the transfers of activity that have taken place across the private/public boundary. It is evidently desirable that the relocation of activity does not in itself lead to a change in the estimate of national output. (2005, p. 36)

However, even when comparisons are possible, output measures need to be interpreted carefully to derive meaningful results on organisational performance.

> When funders and other stakeholders do focus on non-profits' costs, they most often scrutinize cost per output, not cost per outcome … Some funders even provide a fixed amount of funding per output. This focus on cost per output can be counterproductive if it is not married to a focus on maintaining or improving outcomes. (Neuhoff and Searle 2008, p. 36)

Hence, when interpreting the implications of output-based measures, it is important to consider whether the result is consistent with improved effectiveness in achieving outcomes.

Outcomes

Outcomes are defined as the intended results of NFP activities. They are the basis on which the success, or otherwise, of these activities is usually judged. While output measures may attempt to place a dollar value on NFP activities, outcome measures go a step further by trying to capture the full benefits to the recipient over and above the market price or proxy for market price.

According to a guide prepared for NFPs in the United States:

> Outcomes are benefits or changes for individuals or populations during or after participating in program activities. They are influenced by a program's outputs. Outcomes may relate to behaviour, skills, knowledge, attitudes, values, condition, or other attributes. They are what participants know, think, or can do; or how they behave; or what their condition is, that is different following the program. (United Way of America, 1996, p. 2).

From the organisational perspective, outcome measures provide information on how well it is achieving its mission for clients and members. Examples of activities and possible outcomes are provided in appendix B.

In the proposed framework, each of the output categories has an associated class of directly related outcomes, although they may indirectly achieve outcomes in other categories. These are:

- service outcomes in the target communities or individuals (for example, sustained employment placement, improvements in health and fitness of specific groups, creative fulfilment indicators (self-reported survey), or changes in stress related health measures)

- connection outcomes (for example, changes in social capital indicators)
- influence outcomes (possible indicators relate to changes in the allocation of resources, policy changes resulting from advocacy work, or measures of community attitudes)
- existence outcomes (for example, the effect of maintaining or adding to the stock of assets on heritage preservation, cultural identity, freedom of creative expression, or intergenerational equity).

While some outcomes may arise directly from the outputs of an NFP, most are the result of a combination of factors (for example, social welfare outcomes may be the result of NFP and government outputs). This gives rise to the issue of attribution and the difficulties associated with establishing, firstly, the extent to which the activity has been responsible for the outcome (as opposed to environmental or other factors) and secondly, identifying the unique contribution of NFPs.

One approach is to classify the NFP outputs as either:

- sufficient for the outcome to arise
- necessary but alone not sufficient
- enhancing — where the quality or extent of the outcome is greater relative to what otherwise would have happened.

For example, the outcomes of a drug rehabilitation program could be measured by the average number of drug free days achieved in the year following the program. This outcome may be attributable entirely to involvement in an organisation's program (sufficient), in combination with the individuals concerned accessing other services (necessary), or may raise the average number of drug free days achieved (enhancing).

Ted Flack noted that the trend towards clustering of services complicates the task of evaluating just one, suggesting that some need to be evaluated as a package:

> As a consequence of this bundling of services into a whole-of-person or whole-of-family set of integrated services, the measurement of inputs, outputs and outcomes of any one of the elements of the specified services is unlikely to capture the true determinates of service performance. (sub. 29, p. 9)

In addition to attribution issues, outcome measures by themselves rarely provide useful information on how delivery can be improved. This requires examining input and output data and evaluating the process by which the activity is delivered. 'Thus, outcome measurement is an addition to existing data collection efforts, not an alternative' (United Way of America 1996).

Measuring outcomes can also be expensive, and require considerable expertise. Tools that assess the participant before and after undertaking the activity can be viewed as invasive and the cooperation of service users might not always be forthcoming.

Acknowledging these challenges, Moore cautioned against:

> … relying only on *outcomes*. The reason is that while it is extremely valuable to have information about outcomes, the systems that capture reliable information about the outcomes of nonprofit efforts are usually not particularly helpful in managing organizations in the short run. The efforts to measure outcomes are too expensive and too slow to provide comprehensive, fast feedback about how an organization is performing. It is important to measure performance with respect to outcomes, of course. How else could an organisation know if was achieving its ultimate goals. But it would be wrong … to limit performance measurement to outcomes, because that robs nonprofit managers and overseers of the information they need to hold the organization accountable on a real time basis. Nonprofit managers are probably going to need a mix of outcome, output, process and input measures to allow them to recognize value in what they are doing, and find ways to improve their performance. [Emphasis in original] (2003, p. 12)

The problem is magnified when attempting to derive sector-wide measures. Like outputs, outcome measures are not typically expressed in a common metric. Moreover, the differences between the measurement approaches mean that outcome measures from different organisations are not comparable and therefore incapable of aggregation unless converted to a common 'dollar' metric through 'willingness to pay' valuation approaches.

Properly interpreted, outcome measures can be combined with measures of outputs and inputs to guide improvements in NFP efficiency (appendix B).

Impacts

Impacts are the long-term effects produced by an activity, directly or indirectly, intended or unintended, positive and negative (OECD 2007). Impacts are distinguished from outcomes as they capture the longer-term and feedback effects, and spillovers from activities. For example, the impact of improvements in a target community's employment outcomes could be identified in higher household consumption, improved personal security for family members and others in the area, improved health outcomes for the individuals and their households, and the personal value to the individuals of making a worthwhile contribution.

A categorisation of different types of impacts is useful to assist in developing common measures of impacts. The recent literature on measuring community and individual wellbeing (box 3.3) provides some options.

While impact is the most important level in terms of contribution, it is also the most difficult to measure with any degree of confidence as:

- understanding the underlying causal relationship between inputs, outputs, outcomes and ultimate impacts requires evaluating and testing (potentially complex) behavioural models
- data is often partial in nature or difficult to obtain
- some activities pose specific measurement challenges (for example, the spiritual value of retaining connection to land for Indigenous Australians that forms part of their ‘sense of self’).

Structural data

Collection of ‘structural’ data about NFPs (such as size, location, purpose, activity) and their clients or potential clients (such as demographic or labour force characteristics) enables better understanding of both the demand and supply side of the sector’s activities. Information on these attributes is important not only for measuring current contributions but also for monitoring trends in the efficiency and effectiveness of organisations within the sector.

Relevant data includes organisational structures, locations, size, activities undertaken, characteristics of labour force and funding sources. Client data includes information on the clients and members of NFPs, encompassing attributes such as demographic characteristics, income and labour market status. Structural data should be analysed in association with measures derived from the framework to provide a fuller understanding of contributions. There are existing statistical frameworks that support the structural data, much provided by the ABS and AIHW.

Box 3.3 **Approaches to measuring impact**

Wellbeing frameworks

A number of determinants of wellbeing have been identified and these provide a useful basis on which some of the likely impacts of NFPs can be assessed. Based on a review of the literature, and of existing wellbeing frameworks, six broad categories (or domains) of wellbeing can be identified.

1. *Sense of self* reflects outcomes in more than physical and psychological health, picking up spirituality, beliefs and sense of identity. It is the product both of genetics and the environment that a person has grown up in, as well as more recent events in their lives.
2. *Engagement in meaningful activity* suggests that enhanced wellbeing can be gained from paid work, voluntary work, caring and from creative endeavours.
3. *Consumption* of goods and services is usually taken as an indicator of wellbeing. There is strong evidence that access to and enjoyment of a wide range of goods and services enhances wellbeing.
4. *Connectedness to others* recognises that people are social beings and value belonging to social groups, from family to friends to wider connection at the local, national and global community levels. These networks of relationship and community are a source of relational wealth, enriching people's lives through satisfaction of the basic psychological drive for love, acceptance, identity and companionship.
5. The *ability to exert influence* reflects the importance of empowerment of individuals over their own choices, and their participation when it comes to decisions at the household, local and wider levels.
6. *Safety from personal harm* extends beyond physical safety to freedom from persecution, discrimination and other sources of emotional, psychological, financial and material harm.

Measures of Australia's Progress

In April 2006, the ABS published the third issue of *Measures of Australia's Progress* (ABS 2006c). The ABS publishes a summary of the headline indicators on its website annually.

The publication presents indicators across three domains of progress — economic, social and environmental. Each indicator signals recent progress, typically denoting developments over the past 10 years to help Australians address the question, 'Has life in our country got better, especially during the past decade?'. The framework includes both headline and supplementary indicators, and focuses on outcomes rather than inputs or processes. The publication includes special articles that relate to, rather than measure, progress — for example, a feature essay on *Life satisfaction and measures of progress*.

Box 3.3 **(continued)**

International approaches

OECD

The *OECD Factbook* provides a global overview of major economic, social and environmental indicators. Data are provided for all OECD member countries and for selected non-member economies. The information is outcome focused, and is not linked to specific service delivery agencies. The 2008 Factbook's special focus was on productivity — how efficiently production inputs, such as labour and capital, are being used.

New Zealand

The New Zealand Ministry of Social Development produces an annual *Social Report,* which provides information on the health and wellbeing of New Zealand society. Indicators are used to measure levels of wellbeing, to monitor trends over time, and to make comparisons with other countries. A website provides data for social report indicators by regional council and territorial authority areas. The Social Report covers nine 'domains' — unlike this Report, these domains do not directly reflect specific service areas (although there is sometimes a broad connection). A limited number of high level indicators are presented for each domain, but there is no attempt to comprehensively address the full range of objectives of any specific government service.

Commission on the Measurement of Economic Performance and Social Progress

The CMEPSP was established by the French government in 2008. Chaired by Joseph Stiglitz, the Commission included a number of eminent economists including Amartya Sen and Kenneth Arrow. The purpose of the Commission was to consider current approaches to measuring economic and social progress (particularly GDP) and to consider how improvements could be made.

The CMEPSP concluded that the measurement of wellbeing is multi dimensional and, in principle, all dimensions should be considered simultaneously. The dimensions are:

- material living standards (income, consumption, wealth)
- health
- education
- personal activities including work
- political voice and governance
- social connections and relationships
- environment (present and future conditions)
- insecurity, of an economic as well as a physical nature.

Sources: Ministry of Social Development (2007); OECD (2008); SCRGSP (2008); CMEPSP (2009).

3.4 Techniques for valuing the contribution

The proposed measurement framework supports evaluation as it reports on NFP activities from the resources used (inputs) through to the ultimate effects on individuals and the community (impacts). It is, however, a reporting framework rather than a diagnostic one, that is, it does not seek to describe the mechanisms by which inputs are converted to outputs, outputs to outcomes and outcomes to impacts. Two possible conceptualisations of these mechanisms are provided in box 3.4.

Box 3.4 **Mechanisms that underpin the results**

The mechanisms that underpin why NFPs take the approaches they take are based on evidence, theory or a faith-based understanding of cause and effect. Mapping from outputs to outcomes requires articulating how the output acts to achieve the outcome. The VCOSS submission provides a categorisation of the main mechanisms through which these results are achieved, building on Sen’s capability and opportunity paradigm. They describe these as ‘change levers’ for individuals:

- Aspiration — life goals and aspirations and belief in their capacity to influence or control their future
- Capacity — underlying capacity to engage in community, learning or work based on factors such as health, housing and home stability, transport, family issues
- Capability — underlying skill base and support network, affecting their ability to engage in community, learning or work and to influence decisions affecting them
- Opportunity — to participate in community, learning or work and to influence decisions affecting them
- Context — the community or regulatory context in which the participant lives and the effect that has on the above factors.

A similar model for change developed for natural resource management is:

- Understanding — is there an awareness of a problem and of solutions?
- Motivation — is the benefit for the individuals who have to undertake a change greater than the costs that they will have to bear, including intrinsic costs and benefits?
- Resources — are adequate skills, funds and other resources available? Are there constraints in terms of resources or the broader environment?

Change may require action on all fronts or only one if that poses a constraining factor.

Sources: Adapted from VCOSS (sub. DR276); National Land and Water Resources Audit (2002).

While following a project logic (in that the categories of outputs flow fairly naturally into the categories of outcomes), the measurement framework does not map these directly one to one. This recognises that often outputs across a number of categories are needed to achieve outcomes in just one of the categories. Further, outcomes in one category may influence impacts in a number of the categories, for example, volunteering generates not only services for others, but impacts in terms of connections and engagement in meaningful activity for the volunteer. The purpose of evaluation is to map these links and test whether the presumed links have been realised. In many cases evidence is inconclusive, especially where baseline data has not been collected or the counterfactual — what would have happened in the absence of the activity — is difficult to ascertain.

For evaluation to provide a reliable guide, good evaluation design is required. The ‘gold standard’ in medical interventions is a randomised control trial, but this can be difficult to implement with social policies (O’Brien and Bogaards 2009). The development of a collated evidence base on robust links in social policy areas can assist to test assumptions and, as evidence builds, can provide ‘benchmark’ measures. Longitudinal data sets are important to support analysis of the broader trends, but can also be interrogated using econometric methods to isolate some sources of impact from other factors that change over time. Similarly, large unit record data sets can be analysed to control for different life experiences and access to interventions. The information from these types of analyses can provide a guide to the values of outcomes achieved across the dimensions of wellbeing that are supported by available data.

Evaluation is of greatest use where it informs activity design and development, and builds the understanding of the practitioners on what works and what does not. Hence there is value in the process of evaluation as much as in the measures it produces. This ‘micro’ focus measurement of contribution is also where the greatest information will be generated for NFPs and funders to support improvements in efficiency and effectiveness.

There are several different techniques that are used to measure the effectiveness of NFPs and, by corollary, their contribution. Box 3.5 describes several approaches, some which have been designed specifically for use in the sector. These, and other measurement approaches, are discussed further in appendix B. In general, these approaches take a similar approach to measuring inputs and outputs, apply a range of methods for identifying and in some cases measuring outcomes, and vary most in the methods they use to value outcomes. Few go beyond direct effects, but some apply novel approaches to provide indicators of the net benefits associated with NFP activities.

Broadly speaking, the approaches to evaluation fall into two categories — those which draw on the principles of cost-benefit analysis (such as social return on investment) and those which have been adapted from standard accounting methodology (such as social accounting). Reflecting this background, there are significant overlaps between the approaches. For example, the cost-benefit concepts of discounting (to enable benefits to be assessed in current dollar terms) and sensitivity analysis (to check the robustness of results to changes in important assumptions) are also important for determining the social return on investment. Financial proxies, which are elemental to the social return on investment methodology, can also be used in other approaches.

Consultation with stakeholders is an important element of all approaches. All enable the user to consider benefits broadly and to utilise a range of evidence. And all require results to be presented in a transparent way so to be capable of independent verification.

Where these approaches differ is in terms of the specific direction given to the those undertaking the evaluation. For example, the Logical Framework approach provides a template matrix for evaluating performance, and highlights the role of assumptions about the external environment, pointing to what should be monitored to guide continuous improvement and adjustment in programs. Social accounting is neutral on the specific measurement approach used, but emphasises the use of an independent audit to certify the veracity of the information provided (appendix B).

In sum, a number of measurement techniques have been applied to, or developed for, the sector. All are consistent with and all can report within the proposed impact mapping framework. The Commission does not indicate a preference for any technique. They all have their advantages and disadvantages. For example, some methods are seen as being more sophisticated but they may be more demanding in terms of their supporting data requirements. Clearly the quality of measurement results, however, depend not only on the techniques used to produce them but also on the available data and supporting information.

Box 3.5 **Existing approaches to measurement**

Cost-benefit analysis

Cost-benefit analysis provides a comprehensive framework for assessing all the costs and benefits of an activity. It can be used to rank activities according to their net benefits to society. In setting out the principles for using cost-benefit analysis in policy evaluation and decision-making, the Australian Government noted that:

> Its power as an analytical tool rests in two main features:
>
> - costs and benefits are expressed as far as possible in money terms and hence are directly comparable with one another; and
> - costs and benefits are valued in terms of the claims they make on and the gains they provide to the economy as a whole, so the perspective is a 'global' one rather than that of any particular individual or interest group. (2006, p. xi)

Social Return on Investment (SROI)

Social Return on Investment (SROI) was developed to enable NFPs to demonstrate the 'worth' of their activities to potential donors. While embodying the underlying principles of cost-benefit analysis, it uses financial values as proxy indicators for measuring the outcomes of NFP activities.

> … SROI Analysis builds on other approaches to understanding non-financial value by quantifying, and including monetary values of, some indicators of the added value. These are then converted to net present value and divided by the amount of monetary investment to arrive at 'social return on investment.'
>
> …While SROI builds upon the logic of cost benefit analysis, it is different in that it is explicitly designed to inform the practical decision-making of enterprise managers and their investors. (Olsen and Nicholls, 2005, p. 4)

Social accounting and audit

Social accounting and audit attempts to measure the economic, environmental and social contribution of an organisation. After consultation with key stakeholders, the organisation prepares a set of social accounts which sets out the value of the activities undertaken by the organisation during the reporting period. The information gathered, and the measurement techniques used, are determined by the organisation but the accounts are subject to scrutiny by an independent social audit board to ensure that they are a 'fair and honest reflection of what happened during the accounting period' (Robbie and Maxwell 2006, p. 41). Hence, social accounting is a reporting approach:

> … which stresses the need for the identification of socially relevant behaviour, the determination of those to whom the company is accountable for its social performance and the development of appropriate measures and reporting techniques. (Crowther 2000, p. 20)

(Continued next page)

Box 3.5 **(continued)**

Stakeholder Value Management Analysis

Stakeholder Value Management Analysis is a stated preference approach which elicits rankings of organisational activities from key stakeholders. The questions are structured to enable the researcher to infer trade-offs that stakeholders would be prepared to make between activities and hence allow valuation of these activities. This approach was recently utilised by The Smith Family which noted that this approach:

> … has not only enhanced our capacity to measure and enhance the value we are providing to our myriad stakeholders, but has also provided a contemporary evidence base around the stakeholders that provide greatest value to the organisation and that as a consequence demand a greater share of our limited resources … (sub. 59, p. 14)

Results-based accountability (RBA)

Results-based (or outcomes-based) accountability starts with the outcomes the activity or program is intended to achieve. It then works backwards to identify the resources and processes by which those outcomes can be achieved, and the indicators by which success can be measured. RBA attempts to distinguish the specific contribution of the organisations from other factors that can influence outcomes.

> RBA makes a key conceptual distinction between *population accountability* where the aim is to achieve better outcomes for particular groups (such as all children and young people) in a defined geographical area; and *performance accountability* which is intended to improve outcomes for the users of individual services agencies and departments as a contribution towards achieving better outcomes at population level ... Similarly, the distinction RBA makes between 'outcomes' (end results) and process indicators is important, because measuring 'success' on the basis of 'outputs' (which describe service specifications, delivery mechanisms and procedures) alone can be misleading. (UnitingCare Children, Young People and Families, sub. 148, p. 12; emphasis in original)

RBA has a number of characteristics which have made it the preferred approach of a number of government agencies, including some who have mandated its use by organisations which they fund (The Illawarra Forum, sub. 52).

Logical Framework ('log frame')

According to the Australian Government, this approach:

> … is a long established activity design methodology used by a range of major multilateral and bilateral donors, including Australia. It is based on a systematic analysis of the development situation, particularly key development problems, and of the options for addressing those problems. (2005, p. 2)

It involves: establishing the outputs, outcomes and impacts of the activity; outlining the assumptions that link these; and, identifying success indicators. Guidance provided by the Australian Government argues that:

> LFA [Logical Framework Approach] is best started early in activity design. (It is more difficult to use the LFA to review and/or restructure ongoing activities which were not designed using LFA principles and practices). As LFA is an 'aid to thinking', it has widespread and flexible application. (2005, p. 1)