British Journal of Social Work (2025) **55**, 1236–1256
https://doi.org/10.1093/bjsw/bcae186
Advance Access Publication December 11, 2024

Measuring outcomes experienced by users of community services—The development of the community services outcomes tree and the associated survey template

Erin Wilson, Robert Campain*, and Chris Brown

Centre for Social Impact, School of Business, Law and Entrepreneurship, Swinburne University of Technology, Melbourne 3122, Australia

*Corresponding author. Centre for Social Impact Swinburne (CSI), School of Business, Law and Entrepreneurship, Swinburne University of Technology, PO Box 218, Mail H25, Cnr John and Wakefield Streets, Hawthorn VIC 3122, Melbourne, Australia. E-mail: rccampain@swin.edu.au

Abstract

There continues to be growing pressure from human service organizations, funders, and beneficiaries to demonstrate a service's actions are having a demonstrable benefit. This article reports on the method of development and the results of the initial piloting in Australia of an outcomes framework and form of measurement for the human services sector, taking a 'whole of life approach' in recognizing the ways in which life domains interrelate. The design was based on a review of outcomes-related literature, while also using an action research method involving service staff and a consumer advisory body. The result was the Community Services Outcomes Tree (CSOT) and a lean measurement tool in the form of a customizable survey template. We argue the CSOT has passed an initial grounded test in providing meaningful data for each service involved in piloting it and has value in guiding practice and providing a way of measuring and informing outcomes of services.

Keywords: community services; human services; outcomes measurement

BASW
The professional association for social work and social workers
www.basw.co.uk

Accepted: November 2024

Introduction: why outcomes measurement?

The measurement of programme and service outcomes has been steadily increasing and developing in human service organizations both internationally and in Australia (Flynn and Hodgkinson 2001; Schalock and Bonham 2003; MacKeith 2011; Benjamin 2013; Flatau et al., 2015; Bialik 2017; Adams et al., 2020; De Pieri, Chiodo, and Gerli 2023). The sector includes organizations that provide community services (McDonald et al., 2011) such as charitable and not-for-profit organizations, and social enterprises. In Australia, community services organization (CSO) is a label for non-government, non-profit organizations working to deliver a range of services or to represent the interests of a specified constituency. They provide a diverse range of services including health, education, employment, emergency relief, disability services, family services, and housing, among others. Community services comprise small organizations delivering services locally, while some provide a range of services and operate nationally (Casey and Dalton 2006). As part of the piloting for this project, organizations ranged from Uniting Vic Tas that provide a variety of services to a broad cohort of service users, to One Good* Day that provide recovery coaching for people with psychosocial disability, Independence Australia that supports people with disability, and Ganbina who work with Aboriginal and Torres Strait Islander youth and their families with its school to work transition programme. The sector also includes numerous peak organizations that represent member organizations.

Funding agencies (including government) and other stakeholders, such as recipients of services, peak bodies and researchers, are increasingly interested in service delivery outcomes—an understanding of what has changed for service users because of services or initiatives they have received. Human service organizations are expected to report on effective performance of services and engage in performance measurement (McDonald et al., 2011; Pérez Jolles et al., 2016) with the use of this kind of evidence a key strategy for enhancing evidence-based effectiveness in human service organizations (McBeath et al., 2015; Monson, Swords, and Spratt 2022; Flynn, Sundaresan, and Caffrey 2024). Added to this is growing pressure from funders, regulators, and beneficiaries to demonstrate and prove that a service's actions are having demonstrable benefits (DiMaggio 2001; Benjamin 2013; Flatau et al., 2015; Bialik 2017; Mensing 2017; Yang and Northcott, 2019; Adams et al., 2020; Hiruy et al., 2022; De Pieri, Chiodo and Gerli 2023). Research from the UK, the USA, and Canada has shown that the increase in outcomes measurement is largely driven by funder needs to prove a programme's worth (Flatau et al., 2015), with human service organizations increasingly shifting to a 'mission-centred' approach with an emphasis on measuring outcomes (Bialik 2017).

While definitions of outcomes can vary, for the purpose of this article, outcomes have been defined as the 'changes resulting from the implementation of a programme, policy or other intervention' (Ramia et al., 2021: 5) and 'the differences these programs make to the lives of the people they engage with' (Flatau et al., 2015: 4) to prove that service actions are both 'effective and efficient' (Adams et al., 2020: 109).

Outcomes measurement is often attached to meeting certain identified outcomes and these may be based on an outcomes framework. An outcomes framework comprises a set of resources to help stakeholders plan, monitor, and evaluate their activities in an outcome-focused way (Craig 2013; Flynn, Sundaresan, and Caffrey 2024), including a collection of outcomes that can be measured and that are often organized according to key domains. Since the 1990s there has been a plethora of tools, guides and methodologies to support outcomes measurement practice (Benjamin 2013; Adams et al., 2020), with a multitude of frameworks existing which can be confusing for organizations, leading to the fragmentation of outcomes practice (Flatau et al., 2015). These frameworks can provide valuable guidance but the diversity has also been identified as difficult to navigate (Adams et al., 2020). Many have been developed to be applicable to specific initiatives (such as education or employment programmes), they may be designed for a specific sector (such as disability or aged care), while very few have a broad approach that can be applied to all sectors of service provision (e.g. education as well as employment, social inclusion, health, and so on). Alternatively, these frameworks often focus on population-level outcomes (e.g. how many people in a population have certain attributes, such as experiencing homelessness) as opposed to outcomes for individuals.

We sought to address this diversity and fragmentation across the multitude of outcomes frameworks by developing an outcomes framework and measurement method for human service organizations and the community service sector as a whole. Our work was in collaboration with Uniting Vic Tas—an Australian community service organization that works across a variety of service areas (e.g. health, housing, employment, and education). As a result, we undertook a 'whole of life approach' in recognition of the ways in which life domains interrelate. Services—whatever their targeted focus—are often achieving outcomes across this landscape of life domains, yet a narrow outcome approach that focuses on one specific area may not recognize this (Mensing 2017). By outlining a method of outcomes measurement we aim to add further knowledge to this area, recognizing the need for human service organizations to further enhance the value of targeted evidence use (McBeath et al., 2015; Monson, Swords, and Spratt 2022).

The problems and challenges of outcomes measurement as a practice

While there is an identified need to understand outcomes of services and to evidence these in a meaningful and effective way, there are challenges and complexities with undertaking this task. Researchers have noted the difficulties in undertaking outcomes measurement with a body of literature devoted to both outlining and addressing some of the problems (e.g. De Pieri, Chiodo, and Gerli 2023; Flynn, Sundaresan, and Caffrey 2024; Mensing 2017: Stone and Cutcher-Gershenfeld 2001). In brief, these problems include the following.

The challenge of pre-defining outcomes

Effective outcome measurement needs to address the challenge of defining outcomes and identifying metrics that will accurately measure the effect of an intervention on people (Bovaird 2014; Lowe and Wilson 2017; Benjamin, Ebrahim, and Gugerty 2023). However, developing and implementing a prescriptive set of outcomes risks denying or ignoring important service users' progress in life areas that were not initially included in the outcomes framework. There is also the risk that in the choice of outcomes and the related metrics, decision-makers will settle on narrow sets of outcomes that are easy to define and measure (De Pieri, Chiodo, and Gerli 2023), with the organization and its funders having a large say in defining outcomes rather than in partnership with beneficiaries and communities (Benjamin, Ebrahim, and Gugerty 2023).

Being 'client centred' in outcomes measurement

Outcomes measurement typically uses outcomes decided by others and relies on assessments made by clinicians or practitioners. There is a need to align service, organization, and funder outcomes with those identified by service-users and to involve them in the measurement process (Wilson et al., 2013; Mensing 2017; Hagiliassis, Koritsas, and Cuzzillo 2020). Involving service users in defining the outcomes they want to achieve can be empowering and help ensure service provision is relevant (Miller, 2011; Yang and Northcott 2019). As Kroll (2011) states; 'if people are systematically excluded from research participation, their needs, experiences, perspectives are rendered invisible' (p. 67) thereby distorting the evidence base.

Over-simplification of complex issues and quantitative dominance

Outcomes measurement in community services can run the risk of over-simplifying complex issues, reducing and simplifying them through methods that are based on experimental linear designs that frame multifaceted complex issues in mechanistic terms (Craig 2013; Lowe and Wilson 2017; Tse and Warner 2020; Onyura, Mullins, and Hamza 2021), with a simplified linear, predictable 'if-then' causal logic (Bialik 2017; Zappalà 2020). Standardized approaches can overlook cohort features (such as cultural factors) or be inappropriate for some cohorts. Added to this is the risk of reliance on quantitative data as opposed to qualitative which potentially offers a richer picture of service user experience and service features, recognizing—as per a constructivist paradigm of inquiry—'the meanings and interpretations people give to their subjective experiences' (Zappalà 2020: 13).

Service attribution or contribution

Accurately determining the attribution of a specific service or intervention to a particular outcome is one of the most problematic elements of outcomes measurement (Miller, 2011; Bovaird 2014; Bialik 2017; Lowe and Wilson 2017). Whilst decision-makers might be desirous of simplified data sets that seemingly pinpoint effective practice and direct causality, such clarity is elusive. Mayne (2012) in his work on 'contribution analysis' argues that certainty in outcomes measurement is not possible, arguing that *contribution* recognizes the multiple factors influencing a result and considers whether the intervention made a 'noticeable contribution to an observed result and in what way?' (p. 273). Causality is usually of the probabilistic form that the intervention is most likely to have made a difference: 'The aim is to get what Hendricks (1996) calls 'plausible association': whether a reasonable person would agree from the evidence and argument that the programme has made an important contribution to the observed result' (Mayne 2011: 62).

The timing, accuracy, and appropriateness of data collection in community services

In certain circumstances (e.g. crisis), the imposition of outcomes instruments may be potentially offensive, difficult or traumatic for the service participant making respondent burden a key consideration (Quilliam and Wilson 2011; Wilson, Jenkin, and Campain 2011; Yan, Fricker, and Tsai 2019). Also, the 'achievement' of certain outcomes may not easily occur within short-term prescribed timeframes of a funded service window, and what is intended to be measured may not yet be measurable (Ramia et al., 2021).

Lack of resources and staff skills

Community services are often poorly equipped to undertake outcomes measurement making stakeholders resistant to undertake the task. Outcomes measurement demands significant staff time, attention, and funds if it is to yield useful information making it a resource intensive endeavour (Lowe and Wilson 2017; Mensing 2017; Flynn, Sundaresan, and Caffrey 2024). There may be a lack of skilled personnel in undertaking outcomes measurement (Callis, Flatau, and Seivwright 2017; Adams et al., 2020; Ramia et al., 2021; Benjamin, Ebrahim, and Gugerty 2023; Flynn, Sundaresan, and Caffrey 2024), along with a lack of appropriate tools and frameworks to undertake the task (Seivwright et al., 2016).

The above list is not exhaustive but does highlight some of the key concerns with outcomes measurement for community service providers. The task is how to acknowledge and address these concerns while recognizing that there is no one way to undertake outcomes measurement that will satisfactorily address all of these issues. Knight et al. (2017) argue that in undertaking outcomes measurement we need to become comfortable with uncertainty. Outcomes are a result of the system as a whole, not of particular interventions, and there is a need to recognize the reality of 'the messy, complex systems we live and work in' (p. 16). Equally, measurement for the sake of measurement can be harmful to programmes and service user progress particularly if measurement hinders programme delivery (Ramia et al., 2021). Mindful of the need to consider the concerns outlined above, we undertook the development of an outcomes measurement framework and data collection approach that attempted to meet the community services sector need for a method that was straight-forward and meaningful for service users and services, and that provided useful data for a range of stakeholders.

The design logic: the development of the Community Services Outcomes Tree and survey method

We were guided throughout the endeavour by critical thinking that recognizes that human social life and knowledge is often contradictory and contested, and that there are a multitude of perspectives and disagreements over categorizations (Elder and Paul 2008). Critical realism states that the evidence we observe can come close to reality but is always a fallible, social, and subjective account (Bhaskar 1975). Given this, we commenced with four starting assumptions based on our review of literature and previous outcomes work we had undertaken:

1. First, we recognized there is no perfect human services outcomes approach and that, despite extensive work in this field, no stakeholder invested in this field has designed such an approach. We sought to design an approach that was based on pragmatic, critical and realist methodological principles—to recognize the role of subjective information of people in a given context while noting the independent structures that constrain and facilitate people to carry out certain activities (Lawani 2021).
2. There is no easy ‘off the shelf’ fix for community service outcomes measurement. There are hundreds of products, outcomes frameworks, methods and instruments/tools relating to outcomes measurement. This makes applicability for any community service difficult with any approach having limitations. This is even more evident given the variety of services many organizations provide and the array of outcomes across different domains.
3. Considerable resources are needed to collect data, analyse data, and make use of the data. Many organizations are not funded to undertake data collection and do not have the resources to collect data.
4. You can’t do everything. Decisions need to be made when it comes to measuring outcomes including choices about what is most valuable to do in any given context. This means that data will be partial and just one part of understanding the phenomenon in question.

We adopted an action research approach (Oosthuizen 2002; McNiff 2013) as a mechanism of intentional and systematic learning throughout the process. From inception through to piloting, we utilized a cycle of implementing, critical reflection to challenge the findings/knowledge, then making any necessary adjustments. Throughout the process, we maintained notes (journalling) of the strengths and weaknesses across the implementation steps as well as areas to improve/change. We undertook a co-construction process with service users (Monson, Swords, and Spratt 2022) through a representative advisory group to provide advice on the framework and survey design. Similarly, workshops were conducted with staff in two services in different contexts (one regional and one rural) to test the framework and survey approach to ensure that the framework represented what they considered central to service provision outcomes. The evolving framework of domains and outcomes were continually tested for face validity through application to various service contexts, partner organizations and research projects with which we were engaged.

Methods: the development of the Community Services Outcomes Tree (CSOT)

To develop the CSOT, researchers coded 200+ international references which included government, not-for-profit, and academic literature related to outcomes measurement. This collection of documents included outcomes frameworks, data collection instruments, policy documents, quality standards and funding criteria documents. We began the review using purposive sampling to identify approximately twenty outcomes frameworks that spanned more than one community services sector, including the NSW Human Services Outcomes Framework (New South Wales Government, 2017); the Victorian public health and wellbeing outcomes framework (Victorian Government, 2016) and the Outcomes Matrix (Good Finance n.d.), which we coded and thematized according to domains and outcomes (and/or indicators). The method was based on thematic analysis using a predominantly inductive approach whereby researchers identified and mapped themes based on the content of the material. This involves making interpretations of the data content whereby the researcher constructs themes based on an overall consistency and coherence of an emerging framework (Braun and Clarke 2012; Creswell 2009).

We continued to review other relevant literature based on the identification of key domains and outcomes which were used as search terms while also exploring associated references—effectively a snowball technique that continued to reveal potential sources. Many of the identified outcomes frameworks and other relevant indices of outcome attainment were focused on just one domain or life area (e.g. education) or were at the population level (i.e. how many people have certain attributes such as completing year 12 secondary education) which we translated into a community service context for individual outcomes. Throughout the process, we collated the information and continued to confirm and/or change themes until saturation was reached where the outcomes discovered had been previously identified from other sources.

Coding of the data using thematic analysis led to themes and sub-themes or, in this instance, domains and outcomes. We thematically analysed the outcomes under each domain, grouping outcomes with similar themes together and seeking to name the central construct the outcome reflects. In each outcome area, the various sub themes within the outcome are reflected in the outcome descriptions which we developed. A key analytic aim was to identify discrete concepts at outcome level, for example, separating out unique concepts such as ‘safe’, ‘stable’, ‘affordable’ as housing related outcomes. The structure that emerged was:

- Domain (e.g. Housing)

- Outcome (e.g. safe, *or* stable, *or* affordable)
- Descriptor for outcome—based on outcome sub-themes from the literature (e.g. Stable housing—'For example, having secure tenancy or mortgage, being able to maintain tenancy without defaulting, or having long-term housing').

The outcomes framework is presented in the form of a 'tree' with twelve leaves comprising key life areas (Fig. 1).

Each domain comprises a number of discrete outcomes which range in number from 6 (e.g. Health) to 12 (e.g. Family). In total, there are 103 outcomes which includes a 'general' outcome for each domain which combines all the domain outcomes (Wilson et al., 2020).

Throughout the construction of the CSOT we undertook a system of checking by three researchers who worked to ensure consistency across domains, outcomes and descriptions, making decisions about the shape of the framework as it developed. We continually revised for a level of reliability with the reviewed literature and frameworks, while recognizing that the countless outcomes identified cannot be neatly categorized in a way that is precise and unambiguous given the multitude of definitions that overlap and lack in incontestable certainty.

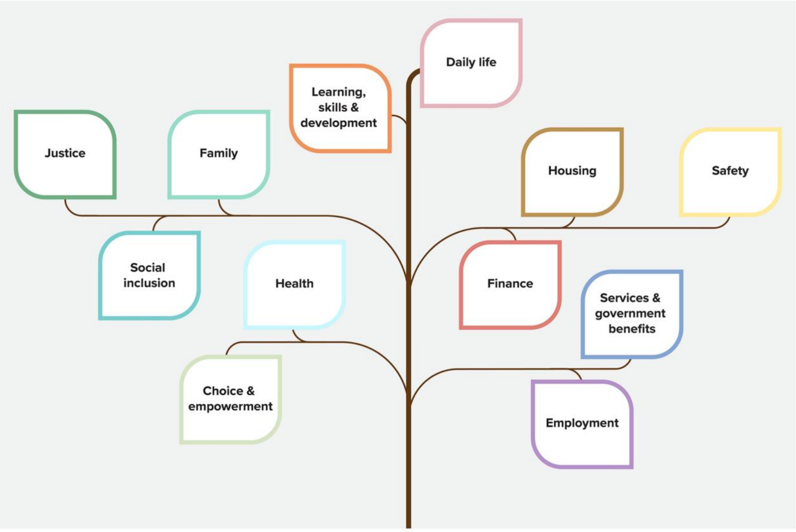

Figure 1. Community services outcomes tree (source: https://communityservicesoutcomestree.com/). Figure 1 is a visual representation of the CSOT which can be accessed via the website. The outcomes framework is in the shape of a tree with twelve major outcome domains represented as leaves. The interactive website enables users to click on each leaf or domain and make visible the nested outcomes within each domain (not shown here).

At its heart, the CSOT framework is an articulation of key outcome concepts relevant to the community services sector. It offers a way to speak about and organize outcomes using a conceptual framework against which to analyse data. The CSOT framework is designed to be 'method-agnostic', that is, it can be used with any data collection method/instrument that measures outcomes that can be aligned to the CSOT domains/outcomes. This enables organizations to maximize all data collected—including across services—by aligning it to a single framework and combining data sources to enable analysis and reporting.

Methods: the development of the data collection survey method

In developing the CSOT, we elected to consider how outcomes could be collected by services that often do not have the resources or skills to undertake outcomes data collection and to enable a method of service user self-report. We sought to address the lack of tools attached to outcomes frameworks by designing a data collection tool in alignment with the CSOT framework.

We adopted a 'minimum data set' approach, using a lean data collection method (Dichter, Adams, and Ebrahim 2016) for each outcome based on service user self-report at a single point in time after service provision. This approach was selected as we noted that services, attempting to use a pre/post approach, had difficulties in collecting initial data at service intake during times of service user crisis, and of securing adequate amounts of matched data at two or more time points. A single time point approach was also a way to ensure that time and resources were not continuously focused on the collection of data thereby impacting staff-service user relationships.

A survey method of data collection was chosen for the opportunity it provides to collect both quantitative and qualitative data, as well as being a viable method for a community organization to administer and collect data from their service users via self-report (Ramia et al., 2021). This was in recognition of the greater validity of this method (as contrasted with proxy report) and as a person-centred approach that values the lived experience of service users and their subjective understandings (MacKeith 2011; Hiruy et al., 2022).

Our survey design approach was based on a method previously used to design an outcomes tool for human services in the disability sector (Quilliam and Wilson 2011). Much of this work was based on the work of Andresen (2000) in her review of a range of outcome tools frequently used in rehabilitation and disability research. Andresen's work identified eleven 'desirable characteristics'. In line with this, we utilized a subset of relevant key criteria for the development of the outcomes measurement tool. Our design response to these desirable characteristics/criteria is available as a supplementary file.

The resultant CSOT survey involves a customizable survey template that outlines the basic survey structure and wording. The survey combines tick-box and open-text responses.

The key focuses of data collection in the survey template involve five key question sets:

1. Outcomes question that seeks to measure the amount of change that has occurred for the service user in relation to the outcome areas. It comprises a list with a rating scale with or without an additional qualitative question seeking to explore the most important area of change. This is informed by Most Significant Change techniques which value respondents reporting on '*why* they consider a particular change to be the most significant one' (Dart and Davies 2003: 138).
2. Contribution question to determine the extent to which the service provider affected the outcome—a list with a rating scale with or without an additional qualitative question to explain what the service did to help achieve outcome/s. We have elected for an extent of *contribution* rather than *attribution* (Mayne 2011; Ramia et al., 2021) noting, however, that the complexity of people's lives and the context in which the service operates makes it difficult to deduce with any certainty that a particular service provision directly led to an identified outcome (Lowe 2013; Lowe and Plimmer 2019). The intent though is to enable a correlation between actions and a desired outcome (Bialik 2017).
3. Barriers to outcomes to identify the systemic, service, and personal barriers. This is useful for advocacy (Bialik 2017) and takes the focus from the individual as being at fault if outcomes are not achieved. This is a tick box list of barriers, with or without an additional qualitative question seeking to explore the most significant barrier to outcomes. The list of barriers is based on an initial review of outcomes and community services literature and has subsequently been enhanced and modified based on recommendations from staff and service users.
4. Service improvement—using a qualitative question/response.
5. Any other comments—a free text response.

Further questions focus on demographics (kept to a minimum based on what is meaningful and useful, and altered based on service specifics), with the potential to add other questions as determined by services and service users.

Methods: piloting the framework and survey

Piloting occurred across a 12-month period in 2022–23. The piloting of the CSOT and survey method in this time involved eleven programmes/services across four distinct service providers, with 293 service users taking part predominantly in two Australian states (Victoria and Tasmania) based on the service provision of the providers. Of these, eight of these programmes/services (with 156 service users) involved Uniting Vic Tas—the Centre for Social Impact (Swinburne) community services partner. Service contexts included: housing/homelessness; early learning; mental health/psychosocial disability; family/parenting; emergency relief; household energy efficiency; education; and employment.

The pilots aimed to test:

1. the ease of the data collection method both for service users and for service staff supporting data collection;
2. the modalities of data collection (to suit different cohorts or contexts);
3. the resources required for data collection;
4. the suitability and meaningfulness of the data collection approach;
5. data analysis approaches and resources required;
6. the utility (usefulness and uses of) reporting;
7. the suitability of the framework as a whole for representing aggregated outcomes across diverse services within a single organization.

Broadly each pilot followed a set of implementation steps:

1. Develop the data collection process:
   Each pilot commenced with a meeting (online or face to face) between key service personnel and the Centre for Social Impact (Swinburne) researchers. This meeting provided information about the CSOT, initial discussion about relevant outcomes, the timeframe in which outcomes occur and the prime timing of data collection, and possible modes of data collection. The survey template was provided by the Centre for Social Impact (Swinburne) researchers, with discussion about how it could be best adapted to suit services and service users.
   Following these discussions, a survey draft was prepared by the researchers and shared with the service for feedback. This involved finalizing the outcomes, customizing the language used in the survey, additional questions, as well as selecting suitable 'barriers' for inclusion in the survey. Finalization of the survey was followed by setting up the chosen data collection mode.
2. Data collection:

Across services the period of data collection differed (ranging from 3 weeks to 8 months) and utilized a range of modalities including: online, hard copy, and assisted completion by staff. In all cases, data collection focused on service user self-report even where this was supported by staff.

3. Data analysis:
   Once data were collected, data were accessed by/shared with the Centre for Social Impact (Swinburne) researchers who undertook analysis.
4. Reporting:
   A brief and accessible Outcomes Report proforma was developed to provide a consistent format to report results to services and stakeholders.
5. Ongoing review
   The Centre for Social Impact (Swinburne) researchers continue to engage with community services as to the utility of the CSOT and measurement methods.

Findings and discussion

The CSOT framework

Overall, the CSOT framework was able to be suitably applied to the range of community services within the pilots. The ‘descriptions’ of each outcome area provided on the website were helpful in clarifying the scope of each outcome, and in supporting staff to further customize the focus of each outcome they included on the survey. During the pilots, only one outcome was identified as not being explicitly captured by the framework: ‘Hope’ or ‘hope for the future’, now incorporated into the outcome ‘Meaning and purpose’.

The visual design of the website and comprehensive capture of outcomes in the Community Services Outcomes Tree provided an accessible and meaningful resource for service staff to engage with. In general, staff readily understood and identified outcomes, and the framework offered the opportunity to focus on a broad array of outcomes beyond the primary ones identified by the funder. However, the in-depth knowledge of the framework held by the Centre for Social Impact (Swinburne) researchers meant that they acted as an informed guide to the framework, able to quickly interpret staff articulations of outcomes into the domains and categories of the CSOT framework. This speeded up the process and provided a level of consistency to outcomes interpretation and selection. This indicates that while the website and visual framework are valuable and capture the breadth of outcomes community services engage in, ‘expert’ navigation and support are helpful. Further piloting without the involvement of researchers would be beneficial in determining the value and ease of CSOT use for service staff.

Customization of language, which includes translation to vocabulary commonly used by the service and service users, was undertaken by service staff about outcomes and more broadly across the survey. This was an important element of ensuring the survey text was relevant to service users. The customization of language and survey content is important for relevance in addressing the unique service context of its users, but this must be managed for ongoing consistency. This mix of customization (flexibility) and consistency is also recommended by other outcomes commentators (Luecking and Weaver, 2022).

Outcomes from services were seen in piloting to overlap with the area of service quality and satisfaction—an area not often included in outcomes measurement (Flynn, Sundaresan, and Caffrey 2024). Service users often connected the concepts of outcome and process or the way the service was delivered—such as being treated with respect, being listened to, or 'heard'. In particular, service users frequently discuss these elements in the question relating to contribution of the service to outcomes. In response to this question, service users often report the way they were treated by particular staff or their experience of the service. Results show that process elements are inextricably bound up with the contribution the service makes to the outcomes. This raises questions about whether process elements should be more explicitly incorporated into the CSOT method and whether, where there are separate data collection processes for outcomes and quality/process measurement, these two data sets need to be re-connected and 'read' together. This connected thinking is consistent with realist evaluation and necessary to make informed judgements about the logic of service design (Pawson and Tilley 1997).

The survey method and data collection

Administration

A variety of survey distribution modes were employed with successful completion across all survey questions regardless of mode. Most services utilized multiple methods of data collection, simultaneously, to suit service users. Modes of data collection included:

Online survey

Eleven services used online surveying including nine services adopting a Qualtrics online survey. Some services felt that online surveying would not be suitable for users or could only be accessible if practitioners provided access to digital devices which was done in several pilots.

Paper based survey

The online survey was able to be converted into a word document and a paper-based survey that service users could complete by hand. This adds a level of resource impost as surveys need to be returned to the service and sent to researchers and entered manually into digital files for analysis.

Practitioner assisted

This method was used in seven of the eleven pilots. While issues of influence might exist when practitioners support service users to provide data, this is offset by the practicality that without this assistance no data are likely to be available. Qualitative responses by service users would seem to indicate they were able to freely express their views, both positive and negative, about the service. However, some caution needs to be applied in assessing results by acknowledging the potential of practitioner influence, both directly and indirectly.

Data value

A variety of responses to outcomes and barriers questions indicated that respondents were able to be discerning in their responses and discriminate amongst a list of items. The contribution question elicited a variety of responses and through qualitative feedback people were able to highlight the role of the service. Often this feedback tended to be more about what people valued about the service rather than specifically focusing on the level of contribution and it may be that the wording of the question needs to be considered. However, the data gained were useful in providing feedback to services about their service provision.

Feedback to date indicates that the data have been valuable and has been used in a variety of ways including as part of formal service evaluations for reporting to funders, as a means of service improvement, for reporting within sectors (e.g. homelessness), for funding applications, and for reporting to senior management as part of wider organization review. The combination of quantitative and qualitative data was able to provide reports that enabled services to have insight into the outcomes of their work, barriers, and feedback on elements of service provision.

Service capacity

Across services, key staff were involved in constructing surveys, administering the data collection method, and in some instances providing support to undertake data collection. While this did take up some staff time, it was not reported as unduly burdensome. Data analysis was more problematic given that it is a skilled function. Skilled data analysts are needed to understand outcomes measurement and analysis of both quantitative and qualitative data, and to oversee the consistency of analysis. Data analysis requires analysis across all data/services, as well as capacity to draw conclusions and generate reports for different uses and for different levels of the organization.

In all cases, data were analysed by a consistent set of Centre for Social Impact (Swinburne) researchers who had substantial expertise in this form of data analysis and in outcomes measurement. Additionally, in one case, the data were also analysed internally by service staff, however, this analysis proved to be inadequate and lacking in understanding of the outcomes logic.

Limitations and further developments

While the piloting to date has identified the value and usefulness of CSOT and the related measurement method, further research is needed to better understand the perspectives of service users and service staff in relation to implementing outcomes measurement. More information is needed about what has worked or not from these perspectives, whether the data capture has adequately represented the outcomes experience of service users, how outcomes measurement can be adequately resourced and importantly, how to maximize the utility of this data at the level of service, organization, and to guide funding.

The mixed modalities of survey completion were beneficial with some service users able to complete the survey independently. However, some required support. While such support is often appreciated when the supporter is reliable and trustworthy, the acquirement of independent self-report requires further development. The establishment of inclusive practices requires further consideration to ensure that all service users have the opportunity to provide feedback on service provision in ways that ensure all service users have an equal opportunity to respond if they wish to do so. This is important in going beyond CSOT piloting to date and ensuring any service-wide evaluation offers equal opportunity to all. This also assists in ensuring data is representative across the cohort.

To date, piloting has been undertaken solely in Australia though we would argue that the CSOT and survey method also has applicability in other Western countries—particularly the UK and USA—based on the international literature which has underpinned its development. Such literature demonstrates that the variety of community services aim for comparable outcomes across similar life domains. Further, the CSOT online access figures show wide viewing particularly in countries such as the USA, China, Germany, Singapore, the UK, France, Canada, and

New Zealand. Further study is required to determine if/how the method is being applied and the results and value of the method internationally—including non-Western countries given website interest from such countries.

Conclusion

Measuring outcomes should serve the purpose of providing information that is valuable to stakeholders and can inform our understanding of service provision and service improvement. A measurement approach needs to be one that is practical and efficient for services to undertake and that enables the views of the recipients of services. We argue the CSOT has passed a grounded test in providing meaningful data for each service involved in piloting it. The CSOT and survey have value in guiding practice and providing a way of measuring and informing outcomes of services, though data analysis requires further consideration as to the type of skilled resource an organization may require. Ultimately, any data collection will always be partial in the story it tells given the complexity of service provision and of service users' lives, and no method can suit all purposes or meet all criteria. The CSOT, though, offers a way of supporting human service organizations to assess their outcomes, and a base for further development and evolution of this undertaking. Further research into the CSOT and measurement tool will be valuable in continuing to refine this endeavour.

Acknowledgements

The authors would like to acknowledge the support of Uniting Vic Tas and thank the service staff and service users in supporting the project from its inception and being involved throughout the various stages of the Consumer Outcomes project. We also acknowledge and thank the staff and service users of One Good* Day, Ganbina, Independence Australia and Moonee Valley City Council. We also thank Mr James Kelly for his technical assistance with the Community Services Outcomes Tree website.

Supplementary data

Supplementary data are available at *British Journal of Social Work* online.

Conflicts of interest. None declared.

Funding

The work was funded as part of an ongoing collaboration between the Centre for Social Impact, Swinburne University of Technology and Uniting Vic Tas.

Ethics approval

All research involving human participants as part this article has been subject to the approval of the Swinburne University of Technology, Human Research Ethics Committee. All participants gave informed consent.

References

Adams, S. et al. (2020) 'Measuring and Accounting for Outcomes in Australian Human Services Charities', *Third Sector Review*, **26**: 108–38.

Andresen, E. (2000) 'Tools of Disability Outcomes Research: Criteria for Assessing the Tools of Disability Outcomes Research', *Archives of Physical Medicine and Rehabilitation*, **81**: S15–S20.

Benjamin, L. M. (2013) 'The Potential of Outcome Measurement for Strengthening Nonprofits' Accountability to Beneficiaries', *Nonprofit and Voluntary Sector Quarterly*, **42**: 1224–44.

Benjamin, L. M., Ebrahim, A., and Gugerty, M. K. (2023) 'Nonprofit Organizations and the Evaluation of Social Impact: A Research Program to Advance Theory and Practice', *Nonprofit and Voluntary Sector Quarterly*, **52**: 313S–52S. https://doi.org/10.1177/08997640221123590

Bhaskar, R. (1975) *A Realist Theory of Science*. Eugene, OR: Harvester.

Bialik, J. (2017) 'Creating a Taxonomy of Outcomes in a Comprehensive Nonprofit Human Services Organization', *Human Service Organizations: Management, Leadership & Governance*, **41**: 201–6. https://doi.org/10.1080/23303131.2017.1321283

Bovaird, T. (2014) 'Attributing Outcomes to Social Policy Interventions – 'Gold Standard' or 'Fool's Gold' in Public Policy and Management', *Social Policy & Administration*, **48**: 1–23.

Braun, V. and Clarke, V. (2012) 'Thematic analysis', in H. Cooper, P. M. Camic, D. L. Long, A. T. Panter et al. (eds.) *APA Handbook of Research Methods in Psychology*, Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. USA: American Psychological Association, pp. 57, 71.

Callis, Z., Flatau, P., and Seivwright, A. (2017) Outcomes Measurement in the Community Sector: Are we Heading in the Right Direction? Bankwest Foundation Social Impact Series, No. 8, Perth, Western Australia, Bankwest Foundation.

Casey, J. and Dalton, B. (2006) 'The Best of Times, the Worst of Times: Community-sector Advocacy in the Age of 'Compacts', *Australian Journal of Political Science*, **41**: 23–38. https://doi.org/10.1080/10361140500507260

Craig, N. (2013) ‘Seeing the Wood and the Trees: Using Outcomes Frameworks to Inform Planning, Monitoring and Evaluation in Public Health’, *Journal of Public Health (Oxford, England)*, **35**: 467–74.

Creswell, J. W. (2009) *Research Design: Qualitative, Quantitative, and Mixed Method Approaches*, 3rd edn. Thousand Oaks, CA: Sage Publication.

Dart, J. and Davies, R. (2003) ‘A Dialogical, Story-Based Evaluation Tool: The Most Significant Change Technique’, *American Journal of Evaluation*, **24**: 137–55. https://doi.org/10.1177/109821400302400202

De Pieri, B., Chiodo, V., and Gerli, F. (2023) ‘Based on Outcomes? Challenges and (Missed) Opportunities of Measuring Social Outcomes in Outcome-based Contracting’, *International Public Management Journal*, **26**: 437–62.

Dichter, S., Adams, T. and Ebrahim, A. (2016) ‘The Power of Lean Data’, *The Stanford Innovation Review*, **14**: 36–41. https://ssir.org/articles/entry/the_power_of_lean_data#

DiMaggio, P. (2001) ‘Measuring the Impact of the Nonprofit Sector on Society Is Probably Impossible but Possibly Useful: A Sociological Perspective’, in P. Flynn and V. A. Hodgkinson (eds) *Measuring the Impact of the Nonprofit Sector*, pp. 249–72. Berlin, Germany: Springer Science and Business Media.

Elder, L. and Paul, R. (2008) ‘Critical Thinking in a World of Accelerating Change and Complexity’, *Social Education*, **72**: 388–91.

Flatau, P. et al. (2015) *Measuring Outcomes for Impact in the Community Sector in Western Australia*. Perth, Western Australia: Bankwest Foundation.

Flynn, S., Sundaresan, S. L., and Caffrey, L. (2024) ‘Putting Outcomes into Practice: The Implementation of a Framework of Outcome Measures within a Child and Family Service’, *The British Journal of Social Work*, **54**: 2378–95. https://doi.org/10.1093/bjsw/bcae037

Flynn, P. and Hodgkinson, V. A. (2001) ‘Measuring the Contributions of the Nonprofit Sector’, in P. Flynn and V. A. Hodgkinson (eds) *Measuring the Impact of the Nonprofit Sector*, pp. 3–16. Berlin, Germany: Springer Science and Business Media.

Good Finance. (n.d) ‘The Outcomes Matrix’, https://www.goodfinance.org.uk/measuring-social-impact/outcomes-matrix

Hagiliassis, N., Koritsas, S., and Cuzzillo, C. (2020) ‘Measuring Goal Realization Associated with Disability Services and Supports: Initial Evidence for a New Tool’, *Journal of Policy and Practice in Intellectual Disabilities*, **17**: 4–12.

Hendricks, M. (1996) *American Evaluation Association Annual Conference*, Atlanta, GA.

Hiruy, K. et al. (2022) ‘Placing People at the Centre of Social Impact Measurement: Current Approaches, Challenges, and Future Direction’, in R. Hazenberg and C. Paterson-Young (eds) *Social Impact Measurement for a Sustainable Future*, pp. 27–48. London: Palgrave Macmillan.

Knight, A. D., Lowe, T., Brossard, M., and Wilson, J. (2017) *A Whole New World: Funding and Commissioning in Complexity*, U.K.: Newcastle University, Business School.

Kroll, T. (2011) ‘Designing Mixed Methods Studies in Health-related Research with People with Disabilities’, *International Journal of Multiple Research Approaches*, **5**: 64–75. http://dx.doi.org/10.5172/mra.2011.5.1.64

Lawani, A. (2021) ‘Critical Realism: What You Should Know and How to Apply it’, *Qualitative Research Journal*, **21**: 320–33. https://doi.org/10.1108/QRJ-08-2020-0101

Lowe, T. (2013) 'New Development: The Paradox of Outcomes-the More We Measure, the Less We Understand', *Public Money & Management*, **33**: 213–6.

Lowe, T. and Plimmer, D. (2019) 'Exploring the New World: Practical Insights for Funding, Commissioning and Managing in Complexity'. Newcastle Business School, Northumbria University and Collaborate CIC, http://wordpress.collaboratei.com/wp-content/uploads/Exploring-the-New-World-Report_Digital-report.pdf.

Lowe, T. and Wilson, R. (2017) 'Playing the Game of Outcomes-based Performance Management. Is Gamesmanship Inevitable? Evidence from Theory and Practice', *Social Policy & Administration*, **51**: 981–1001.

Luecking, A. and Weaver, K. (2022) *Social Sector Hero. How Government and Philanthropy Can Fund for Impact*. Rockville, USA: Clear Impact.

MacKeith, J. (2011) 'The Development of the Outcomes Star: A Participatory Approach to Assessment and Outcome Measurement', *Housing, Care and Support*, **14**: 98–106.

McBeath, B. et al. (2015) 'Organizational and Individual Determinants of Evidence Use by Managers in Public Human Service Organizations', *Human Service Organizations: Management, Leadership & Governance*, **39**: 267–89. https://doi.org/10.1080/23303131.2015.1044588

McDonald, C., Craik, C., Hawkins, L. and Williams, J. (2011) *Professional Practice in Human Service Organisations. A Practical Guide for Human Service Workers*. Australia: Taylor and Francis Ltd.

McNiff, J. (2013) *Action Research: Principles and Practice*, 3rd edn. Abingdon: Routledge.

Mayne, J. (2011) 'Contribution Analysis: Addressing Cause and Effect', in R. Schwartz, K. Forss and M. Marra (eds) *Evaluating the Complex*, pp. 53–96. Piscataway, NJ: Transaction Publishers.

Mayne, J. (2012) 'Contribution Analysis: Coming of Age', *Evaluation*, **18**: 270–80.

Mensing, J. F. (2017) 'The Challenges of Defining and Measuring Outcomes in Nonprofit Human Service Organizations', *Human Service Organizations: Management, Leadership & Governance*, **41**: 207–12.

Miller, E. (2011) Measuring Personal Outcomes: Challenges and Strategies. Institute for Research and Innovation in Social Sciences.

Monson, T., Swords, L., and Spratt, T. (2022) 'Establishing Outcome Measures in Practice: Developing a Model for Services Working Therapeutically with Children and Families', *The British Journal of Social Work*, **52**: 3501–21.

New South Wales Government (2017) *Finance, Services & Innovation, Human Services Outcomes Framework Guide*. New South Wales Government, Finance, Services & Innovation.

Oosthuizen, M. J. H. (2002) 'Action Research', in K. Williamson (ed.) *Research Methods for Students, Academics and Professionals. Information Management and Systems*, 2nd edn. pp. 159–175. Bathurst, Australia: Charles Sturt University.

Onyura, B., Mullins, H., and Hamza, D. M. (2021) 'Five Ways to Get a Grip on the Shortcomings of Logic Models in Program Evaluation', *Canadian Medical Education Journal*, **12**: 96–9. https://doi.org/10.36834/cmej.71966

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*, 1st edn. Thousand Oaks, CA: Sage.

Pérez Jolles, M. et al. (2016) 'Factors Associated With Managerial Innovation in Public Human Service Organizations', *Human Service Organizations: Management, Leadership & Governance*, **40**: 421–34. https://doi.org/10.1080/23303131.2016.1184208

Quilliam, C. and Wilson, E. (2011) 'Literature Review—Outcomes Measurement in Disability Services: A Review of Policy Contexts, Measurement Approaches and Selected Measurement Tools', Deakin University, Burwood, https://hdl.handle.net/10536/DRO/DU:30042076

Ramia, I. et al. (2021) *Roadmap to Outcomes Measurement, Your Step-by-Step Guide to Planning, Measuring and Communicating Social Impact*. UNSW Sydney: Centre for Social Impact.

Schalock, R. L. and Bonham, G. S. (2003) 'Measuring Outcomes and Managing for Results', *Evaluation and Program Planning*, **26**: 229–35.

Seivwright, A., Flatau, P., Adams, S., and Stokes, C. (2016) *The future of outcomes measurement in the community sector: Social Impact Series, No. 6*, Bankwest Foundation.

Stone, M. M. and Cutcher-Gershenfeld, S. (2001) 'Challenges of Measuring Performance in Nonprofit Organizations', in P. Flynn and V. A. Hodgkinson (eds) *Measuring the Impact of the Nonprofit Sector*, pp. 33–57. Berlin, Germany: Springer Science and Business Media.

Tse, A. E. and Warner, M. E. (2020) 'The Razor's Edge: Social Impact Bonds and the Financialization of Early Childhood Services', *Journal of Urban Affairs*, **42**: 816–32.

Victorian Government (2016) Dept. of Health and Human Services, *Victorian public health and wellbeing outcomes framework*, Victorian Government.

Wilson, E., Jenkin, E., and Campain, R. (2011) *Outcome Measurement of Community Based Mental Health Services in Western Australia: Literature and Concept Summary*. Melbourne: Inclusion Matters.

Wilson, E. et al. (2013) 'An Accessible Survey Method: Increasing the Participation of People with a Disability in Large Sample Social Research', *Telecommunications Journal of Australia*, **63**:24–1.24.13.

Wilson, E. et al. (2020) 'The Community Services Outcomes Tree', Swinburne University of Technology, Hawthorn, https://communityservicesoutcomestree.com/.

Yan, T., Fricker, S., and Tsai, S. (2019) 'Response Burden: What Is It and What Predicts It', in P. Beatty, D. Collins, L. Kaye, J.L. Padilla, G. Willis and A. Wilmot (eds) *Advances in Questionnaire Design, Development, Evaluation and Testing*, pp. 987–1017. Hoboken, NJ: Wiley & Sons Inc. https://doi.org/10.1002/9781119263685.fmatter

Yang, C. and Northcott, D. (2019) 'Together We Measure: Improving Public Service Outcomes Via the Co-production of Performance Measurement', *Public Money & Management*, **39**: 253–61. https://doi.org/10.1080/09540962.2019.1592906

Zappalà, G. (2020) 'Meaningful Evaluation: A Holistic and Systemic Approach to Understanding and Assessing Outcomes', *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, **12**: 1–19.

British Journal of Social Work, 2025, 55, 1236–1256
https://doi.org/10.1093/bjsw/bcae186
Original article