



Elizabeth-Rose Ahearn* and Cameron Parsell

The Inspection of Marketised Models: Audit, Evaluation, and Service Beneficiaries

<https://doi.org/10.1515/npf-2024-0047>

Received July 22, 2024; accepted February 26, 2025

Abstract: Inspection is an institutionalised feature of social services that aims to ensure accountability and improve quality and outcomes for recipients. Successive changes in administration and funding have created a hybridised market of public, private, and third-sector providers, leading to a complex institutional environment. Our study seeks to understand how the chosen modality of inspection in different social services is influenced by this environment. By combining public statements of inspection from social service organisations with qualitative interviews with sector leaders and experts, we reveal a stark divide: audited services operating within quasi-markets, on the one hand and evaluated services relying on block grant-funding, on the other. These differences are underpinned by embedded expectations of whether services should change outcomes or simply provide care, in turn influencing the design, administration, and quality of services.

Keywords: social services; marketisation; evaluation; audit; institutional theory

1 Background

The purpose of this article is to investigate how characteristics of social services influence the modality of inspection they undergo. *Social services* encompass various operating models and missions, such as education and training, counselling and case management, provision of welfare and material aid, care, and community development (Considine 2022; Tan and Harvey 2016). Previously offered under public assistance, neo-liberal reforms have now shifted these responsibilities to a mixed economy comprising private, public, and third sector providers (Barman 2007; Considine 2003; Salamon 1987).

The inspection of social service quality and performance is a common expectation among all providers (De Waele et al. 2021). Broadly, it aims to meet two key

*Corresponding author: Elizabeth-Rose Ahearn, School of Social Science, The University of Queensland, Brisbane, QLD, Australia, E-mail: er.ahearn@uq.edu.au. <https://orcid.org/0000-0002-5739-3269>

Cameron Parsell, School of Social Science, The University of Queensland, Brisbane, QLD, Australia. <https://orcid.org/0000-0002-7289-896X>

rationales: accountability and the establishment of organisational legitimacy through data and evidence (Deegan 2002; Gray, Owen, and Adams 2009; Moxham 2014), and improvement of service quality and outcomes through understanding causal mechanisms and consumer experiences (Alkin and Taut 2002; Cordery and Sinclair 2013; Reinertsen Bjørkdahl, and McNeill 2022; Weiss 1999). A range of different modalities are employed to uncover the data and evidence required including social audit, evaluation, social impact assessment, and certification (Ahearn and Mai 2023; Dahler-Larsen 2011). Underpinning these modalities is a strong rationalist justification that if we can measure and understand service performance we can then act to bring about improvements, or at least mitigate risks (Hillman et al. 2013). However, adopting an institutional perspective we recognise that “organisational behaviour occurs through and is a consequence of taken-for-granted beliefs, schemas, and values that originate in a larger institutional context” rather than direct “primarily responses to market pressures and efficiency dynamics” (Greenwood 2008, p. 433).

Although the modalities described above “are not carried out in the same contexts and not by the same people... from a sociohistorical perspective, it is also reasonable and analytically beneficial to consider [them] under [a] common rubric” (Dahler-Larsen 2011, p. 11). The key analytic benefit of recognising the common features of these practices of measurement and inspection is to show how they have become institutionalised. They now form part of a protected discourse that is considered “virtually sacred, about which the dominant forces in society do not pose questions” (Dahler-Larsen 2011, p. 3). Although Dahler-Larsen (2011) adopts the term *evaluation* to describe this common rubric, our analysis here requires the differentiation between evaluation and audit. Therefore, we describe this common rubric as *inspection*.

The institutionalisation of inspection throughout society, particularly in social services, arises from mutually reinforcing discourses that aim to establish the legitimacy of an organisation or its undertaking (DiMaggio and Powell 1983; Phillips et al. 2004). Actors establish their legitimacy by emphasising the inspections they have undergone, thereby limiting the future representations of legitimacy available to them and other organisations in their environment (DiMaggio and Powell 1983; Phillips, Lawrence, and Hardy 2004). This occurs within public discourse such as annual reports, websites, marketing, and promotional material (Fairclough 1992). In turn, as organisations in the sector “respond to an environment that consists of other organisations responding to their environment, which consists of organisations responding to an environment of organisations’ responses” (DiMaggio and Powell 1983, p. 149), actors are left with no choice but to align with inspection.

While such institutionalisation renders inspection as mandatory, it is not always seen as a fruitful activity, so actors may attempt to avoid, or only ceremonially adopt

inspection (Dahler-Larsen 2011; Dhanani and Connolly 2012; Meyer and Rowan 1977). This is particularly true for organisations that through the pronounced adoption of legitimised values, norms, ceremony, and myths are granted institutionalised status (Dahler-Larsen 2011; Meyer and Rowan 1977). Meyer and Rowan (1977) illustrated that institutionalised organisations avoid inspection as acknowledging the need for inspection “accompanies and produces illegitimacy” by undermining the assumption of their inherent legitimacy.

From the torrent of scholarship and attention focused on how to inspect social services, it is clear that this is not a fully institutionalised field. But why not? The core values underpinning social services of charity, altruism, and care are fundamental aspects of human nature, and ascribed values throughout most societies (Kurzban, Burton-Chellew, and West 2015). They are taught and embedded into other institutions such as religion, the family, and democracy. And yet, we demand heavy scrutiny of them in their organised form. Therefore, this study aims to generate a deeper understanding of how the institutionalised environment surrounding social services generates this pressure to conduct inspection. To achieve this, we first present an overview of literature relating to social service administrative reforms to provide insight into the study context. Next, we draw on two related data sources: a content analysis of public descriptions of inspections undertaken in Australian social services, and qualitative interviews with social sector experts and leaders.

2 Literature Review

From the inception of professionalised charity, inspection has consistently emerged as a fundamental feature. Inspection was first directed towards the recipients of charity (Katz 1986). Charity organisations and their volunteers were expected to investigate recipients to differentiate between the deserving and undeserving poor (Barman 2007; Katz 1986; O'Brien 2015). Broadly, the deserving poor were those who could not be faulted for the position they found themselves in, while the undeserving poor were those who, through indolence or moral failure, were considered responsible for their own suffering (Katz 1986). Assessment and centralised recording of individuals deemed underserving of charity represented the first formalised measurement of social services (Barman 2007; Clarke and Parsell 2022; Powell 2007; Parsell, Clarke, and Perales 2022).

At the beginning of the 20th century, social discourse began to question the role of charitable organisations providing social services, leading to the emergence of state funded and administered social services (Barman 2007). Charitable organisations rebutted this trend by using measurement to quantify the level of need, thereby legitimising their place as social service providers (Barman 2007; O'Brien 2015).

Nevertheless, the rise of the welfare state was seen across many industrialised countries whereby the state took responsibility for providing key social services (Katz 1986; Parsell, Clarke, and Perales 2022). Charitable organisations continued to play various roles, either as contractors to the state or with volunteers providing services alongside the state to address its limitations (Parsell, Clarke, and Perales 2022; Quadagno 1987; Young 2002). There was a brief period when social services went unmeasured (Barman 2007). However, concerns about the quality of state services such as schools and medical facilities led to the introduction of a new age of science-driven assessment, also known as evaluation (Alkin 2004; Patton 2008; Vedung 2010). This application of social science research, influenced by medical and engineering models of causal attribution, examined whether and how social services could change individual behaviour and social conditions (Alkin 2004; Campbell 1979; Vedung 2010). Subsequent evaluation waves have shifted the focus toward different stakeholder groups, yet collecting outcomes-based data from program recipients remains a core aspect of modern evaluation practice (Ahearn and Parsell 2024).

In a parallel evolution, the social accounting and audit movement developed as a necessary extension of conventional accounting (Cooper and Sherer 1984; Williams 1987). While the conventional movement focuses on shareholder interests and financial transparency, in response to corporate scandals and gaining momentum of the corporate social responsibility in the 1970s the social accounting movement began (Gibbon and Dey 2011; Gray et al. 1997, 2009; Lingane and Olsen 2004). Taking a broader focus on organisational accountability, social accounting and audit contends that “organisations have a duty to discharge information pertaining to their social and environmental interactions to a wider group of constituents than simply financial stakeholders” (Spence 2009, p. 206). Initially focused on corporate social impact (Epstein, Flamholtz, and McDonough 1976; Ramanathan 1976), the field has since developed a suite of approaches for providing an account of public and charitable sector performance, alongside audit-based quality, practice, and delivery standards (Cordery and Sinclair 2013; De Waele et al. 2021; Johnsson et al. 2021). Despite this adjusted focus on service and product delivery standards, social accounting retains its original market-aligned emphasis on providing instrumental information to stakeholders through the examination of business processes (Gray 2001). Like evaluation, social accounting and audit have become instrumental tools for facilitating the tenets of subsequent public sector neo-liberal reforms.

2.1 Inspection and Neo-Liberal Reforms

Measurement, including evaluation and audit, continues to provide a means to systematically gather information about a social service to determine its quality, and

inform decision-making (Benjamin, Ebrahim, and Gugerty 2023; Stake 2001; Weiss 1999). However, measurement not only reflects but also constitutes reality and therefore this process introduced new norms in service delivery (Barman 2007). By framing social services as something measurable and quantifiable, inspection facilitated comparison and therefore competition among services, enabling the next stage of social service delivery characterised by neo-liberal reforms (Dahler-Larsen 2011; Lakoff 2014).

These reforms saw private-sector management principles introduced to social services through disaggregating delivery to a new mixed-economy of local governments, charitable and non-profit organisations, and private providers (Christensen and Lægreid 2011; Lapuente and Van de Walle 2020; Salamon 1987). This allowed the state to transfer responsibility and risk away from ministries, while encouraging new actors to innovate and deliver services more efficiently (Considine 2001; Lapuente and Van de Walle 2020). Measurement became central in these marketised arrangements through enabling accountability reporting, which in turn facilitates competition and incentivisation (Bovens 2007; Cordery and Sinclair 2013; de Boer 2023; Eikenberry and Kluver 2004; Lapuente and Van de Walle 2020).

2.2 Modality of Inspection

In Australia, like many other states from Europe, North America, and the UK, these reforms have led to a hybrid social sector where public, private, and third-sector entities operate social service financing and delivery (Onyx, Cham, and Dalton 2016; Salamon 1987). While the government remains the key funder of social services (Australian Charities and Not-for-Profits Commission 2023), competitive tendering and marketised models have entrenched the values of innovation and competition into the social services environment (Considine 2003; Dees 2012). However, without true market forces, performance measurement becomes a central requirement to facilitate accountability and demonstrate adherence to these values (Kendall and Knapp 2000; Sawhill and Williamson 2001).

A review of performance measurement practices across different hybrid social sector models identified four key types of data used (Ahearn and Mai 2023). Financial data addresses the longstanding concern of all organisations to avoid fraud or misappropriation. Perceptual data, used in corporate social responsibility and investing models, aims to measure board and investor satisfaction with the organisation's impact. Compliance data encompasses rating-based measures of quality, certification, and social audits, ensuring that appropriate procedures and policies are in place to protect financial, governance, and delivery standards. Finally, effectiveness data, stemming from outcomes measurement and evaluation. The review

highlighted that these different modalities perpetuate the values and logics present within the originating institutional environments (Ahearn and Mai 2023). As a result, epistemological differences in whose perspectives on the performance of social entities is most valued is embedded within the data collection processes. For those originating within the market environment, there is an emphasis on upward accountability to the information needs of funders and powerful stakeholders (Ebrahim 2005). In contrast, data following the institution of empiricism focuses on causal change and service recipient experiences (Ahearn and Mai 2023). This highlights the critical risk of social services being measured and managed to satisfy upward rather than downward accountability requirements, thereby threatening the underlying value and logics expected of social service (Ahearn and Mai 2023; Ebrahim 2019).

These four modalities of data are typical of the common rubric identified by Dahler-Larsen (2011) and referred to here as inspection. Despite deriving from different sociopolitical origins, following different procedures, and having different consequences, they all represent the institutionalisation of measurement. Recognising this institutionalisation is key to informing our analysis, but the differences between modalities can offer equally constructive insights. Specifically, the presence of different modalities of inspection reflects the complex institutional environment in which social services now operate (Greenwood et al. 2011). As the institutions of the empiricism, the state, and the market have entered the space once solely occupied by charity, new logics, values and myths have been adopted (Dees 2012; Greenwood et al. 2011). These changes bring with them different expectations about what can and should be inspected. Therefore, in this study, we scrutinise the modalities of inspection and examine how social service actors make sense of these modalities, how they are produced, and how meaning is constructed through the institutionalised environment. We achieve this by posing the following research questions.

Therefore, in this study, we scrutinise the modalities of inspection and examine how social service actors make sense of these modalities, how they are produced, and how meaning is constructed through the institutionalised environment.

3 Methods

The study combines two forms of qualitative data: public descriptions of adopted inspection modalities by large Australian social services organisations, and interviews with leaders and experts from the social service sector. The former is directed predominantly at the first research question, and the latter towards the second.

3.1 Naturally Occurring Data

3.1.1 Search and Screening

The first part of our data analysis consists of summarised and codified text published by large not-for-profit social service providers about their core service areas, and the modalities of inspection they use.

Our search sample-frame was identified using the Australian Charities and Not-for-Profit Commission (ACNC) mandatory revenue reporting data for the years 2019, 2017 and 2015. We focused on non-profit organisations that provide social services, therefore excluding universities, philanthropic (grant-making) foundations, and pro-bono legal services. We further focused on organisations that have experienced consistently high income over the preceding 5-year period to ensure a richer sample of organisational activity. High income stability was expected to increase the organisation's ability to invest in long-term sustainability, continuous improvement, and quality processes such those under the banner of inspection (Bach-Mortensen and Montgomery 2018). From the sample-frame, a final group of 50 organisations had remained consistently within the top 100 high-income organisations during the 2013 to 2019 ACNC reporting period.

The website listed on each charity's respective ACNC profile was used to conduct the search. The search terms "evaluation", "audit", "certifi*", "accredit*", "research", and "quality" were entered successively in the website search bar provided and Google Advanced Search within the site domain. We also explored key headings and tabs on the websites, such as 'Publications,' 'Our Research,' 'Our Governance,' and 'Accreditations,' to find text related to inspection. A degree of exploration was required to accomplish this as the layouts and content of each website varied.

Each organisation's self-identified service areas were located and recorded, for example "Emergency Housing", "Financial Wellbeing" and "Aged Care." These were often hosted on pages titled 'Our Services' or 'What we do'.

All references to inspection, such as financial audit statements, certification, or reference to evaluation projects made by each organisation were recorded in an excel file. Further, the text describing each inspection modality was analysed to identify the service area or organisational element it related to. These public statements on inspection modalities constitute what de Boer (2023) describes as "message in a bottle accountability" statements, made to no specific forum but rather a wider, unknown public. These contrast with targeted or principal-agent accountability statements. While the inspection referenced could be resulting from a principal-agent relationship, it represents a separate form of public accountability.

3.1.2 Content Analysis

Conventional content analysis was utilised to identify categories of the service areas described by the organisations (Hsieh and Shannon 2005). There was considerable homogeneity in the naming conventions and descriptions of service areas among different organisations, for example “Aged Care” was exclusively described as “Aged Care”. This analysis identified 19 key service areas (see the x-axis of Figure 1).

The 19 service areas identified were used to conduct a deductive content analysis of the social service targeted by the inspection. For example, the evaluation of a

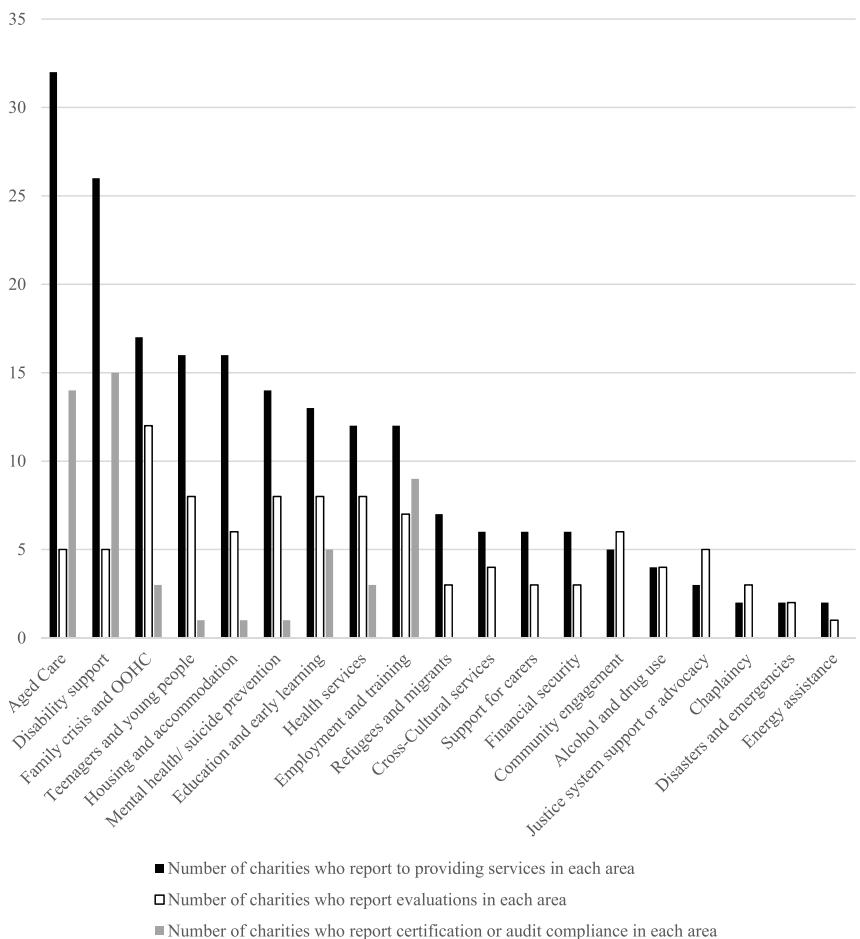


Figure 1: Organisational reporting of modalities of inspection adopted within different social services.

service which focused on “enabling young people at risk of, or experiencing, homelessness to successfully participate in mainstream education, training and employment” was classified as a service for “teenagers and young people”, “housing and accommodation”, “education or early learning”, and “employment and training.” Similarly, accreditation by the Aged Care Quality and Safety Commission was coded as “aged care”, and assessment against the Australian Children’s Education and Care Quality Authority National Quality Standards as “education and early learning.” Inspection that applied to the entire organisation such as Information Security Management System ISO27001, were also recorded, but not coded against service areas.

Finally, the descriptions of inspection were deductively coded against the four modalities described above: financial, perceptual, compliance, and effectiveness. The results of this coding are demonstrated in Figure 1, which shows the number of organisations that claim on their websites to provide and conduct inspection of each service area.

3.2 Qualitative Interviews with Sector Experts and Leaders

3.2.1 Participant Recruitment and Sample

Qualitative interviews were conducted with 20 social service leaders and experts to examine how actors make sense of the selection of modalities towards different social services. To be eligible, participants needed to have considerable experience in the inspection or governance of social services.

A purposive sampling approach was adopted to ensure that interview participants had substantial knowledge regarding the use of evaluation and performance measurement in social services (Armstrong 2010). Initially we searched the staff repositories of large not-for-profit organisations for references to key staff. Examples of job titles we looked for include Evaluation Manager, Continuous Improvement and Learning Director, Chief Executive Officer, or Research Lead. We also utilised the social media platform LinkedIn to identify relevant staff members. Additionally, snowball sampling was employed, where participants were asked to identify other key experts in the field who should be included.

3.2.2 Interviews

The interviews were loosely structured and focused on three key topics relating to decision-making processes underpinning measurement: (1) what gets measured, (2)

the approaches and methods utilised, and (3) the subsequent actions taken based on findings and recommendations.

We adopted the active interview style (Holstein and Gubrium 1995), where respondents are seen not merely as repositories of knowledge but as constructors of knowledge in collaboration with interviewers. In this way, the interview “cultivates meaning-making as much as it ‘prospects’ for information” (Holstein and Gubrium 1995, p. 5). The approach contrasts with the “vessel-of-answers” view of participants, which cautions interviewers to be wary of how they ask questions to avoid biasing the subject’s responses. Instead, we viewed respondents as active participants who constructively add to, take away from, and transform the facts and details they share.

The active approach was important because the interview participants were not only experts in the substantive content of the interviews, but also in the research process itself. To enhance the quality of the information discussed, it was crucial to treat them as collaborators and provide full insight into the study being conducted. All interview participants had held multiple roles and positions within the social and public sectors related to research and evaluation, either as producers or consumers of evidence and data. Respondents actively constructed and repositioned their perspectives and the past, immediate, or future realities from which they drew their responses. Therefore, to fully involve participants in the research process the interviewer shared findings, emerging theories, and examples of the research to create a collaborative environment.

3.2.3 Qualitative Analysis

All interviews were recorded, transcribed, and then interview transcripts were analysed thematically. Each transcript was read in its entirety twice. Next, inductive coding was completed in NVivo 12 to allocate text towards key themes present across the transcripts. By theme, we mean “an implicit topic that organises a group of repeating ideas [which] enables researchers to answer the study question” (Vaismoradi et al. 2016, p. 101). Deductive coding was then conducted to ensure all text relating to the themes were captured. Examples of these themes, and those with the most references pertaining to this analysis, are “service areas”, “funders”, “decision making”, “board governance”, “organisational structure” and “learning.” Finally, all text allocated to these themes was read in a narrative form, in the context of other text relating to that theme, rather than the original interview it arose from. At this stage what we observed was a pattern and sequencing of concepts relating to each other and consistent framing of these relationships. The description of these findings below articulates these themes and draw on key quotes from participants which demonstrate these structures. The quotes presented below are verbatim, with only minor edits for clarity such as the removal of excessive filler words and stutters.

4 Findings

4.1 Naturally Occurring Claims of Inspection

All organisations included in the review published audited financial statements on their websites. No references or indication of perceptual performance measurement could be identified. Neither of these findings are unexpected as all organisations registered with the ACNC must report on financial data, and perceptual measurement is predominantly used in CSR (Ahearn and Mai 2023).

Figure 1 demonstrates the number of organisations who refer to their provision of social service areas, and the collection of compliance or effectiveness data pertaining to these areas. It visually illustrates that while some form of inspection is undertaken in relation to all service areas, there is a clear divergence between the use of audit to determine compliance, versus the use of evaluation to determine effectiveness.

Audit and compliance data was predominantly referenced in relation to disability support, aged care, out-of-home care (OOHC), employment and training, and health services. There was a high degree of consistency between the standards used to assess these, demonstrating the high degree of regulatory oversight in these fields. Respectively, the certifications were the NDIS Commission Practice Standards, Aged Care Quality and Safety Commission Certification, National Standards for Out-of-Home Care, Standards for Registered Training Organisations, and the National Safety and Quality Health Services Audit. Organisations also adopted additional quality standards that did not pertain directly to service areas, particularly the ISO9001 quality management standard.

Effectiveness data took the form of evaluation projects and predominantly related to family crisis services, programs for teenagers and young people, mental health, education and early learning, health services, employment and training, and housing and homelessness services. It is also important to note the service areas which were evaluated despite not being an area the organisation stated to work within on their websites. Specifically, community engagement, alcohol and drug use, justice system support, and chaplaincy were evaluated by charities that did not list them as their focus areas.

Aside for health services, employment and training, and education and early learning, what struck the researchers throughout this search and analysis process was the clear divide between services that conducted audits, and those that conducted evaluation. Generating an empirical understanding of this from actors involved in social service organisations was a core focus of the subsequent qualitative interviews.

4.2 Qualitative Interviews with Evaluation and Social Sector Leaders

4.2.1 Variations in Adopted Modalities of Inspection

The intention of wanting to understand the “behind the scenes” of inspection was expressed to all interview participants. The subsequent detail provided by participants illustrate both definitive explanations and intriguing contradictions and oversights, shedding light on the role of institutional pressures in setting inspection priorities.

As identified in the literature, availability of financial resources to conduct evaluation was a core feature (Bach-Mortensen and Montgomery 2018). However, our interviews revealed that rather than being a barrier to overcome, funding acted as a clear signal directing what would be evaluated. As expressed by one evaluation manager *“if there’s a line item in the funding budget, then we do one”* (Participant 16). However, alongside funder accountability expectations the opportunity for service improvement was also raised. The balance between these two pressures was well articulated by one service quality manager.

Participant 13: So generally, historically the evaluation work has been very project based. There is a, either, extrinsically funded evaluation project, say a funder for a service, says you get this amount of additional dollars additional set aside to do an evaluation of this ... Or intrinsically where, for example, we might have a service that we think may need some additional support to develop a stronger case for its refunding, or we think it’s it has as a new service type that we want to evaluate for our own purposes.

The decision-making process behind conducting an audit illustrated similar factors. Regulatory pressure was particularly strong as many funding models have distinct standards required for compliance as outlined above. In relation to the improvement aspect of audit to ensure compliance this was also apparent, and organisations would adopt additional standards not required by their funders or regulators to go above and beyond the requirements.

These common underlying motivations to conduct inspection concur with existing scholarly literature (Cordery and Sinclair 2013; Reinertsen et al. 2022). Beyond this, a divide between services inspected via audit, and those inspected via evaluation pervaded in the interviews.

Aged care and disability services were often omitted when participants provided an overview of evaluation within their organisation’s service portfolio. For example, a manager at a large organisation which provided many services, including aged care and disability support, stated that evaluation was *“supporting almost all”* of the program areas. They continued, *“so we have the big projects in terms of outcome*

measurement and outcome evaluation in alcohol and other drug services, youth homelessness, family violence, [homelessness service], and financial counselling” (participant 8).

The lack of application of evaluation to aged care and disability services was also directly interrogated in interviews. However, even when prompted, many participants still overlooked these service areas and instead drew examples from other areas, such as housing and homelessness or financial support. Other participants translated this to a lack of evaluation readiness in disability and aged care services. Specifically, they drew on challenges in staff culture, training, and the availability of measurement tools. For example, the following participant was a senior manager at a charity with a whole of organisation impact framework.

Interviewer: Do you provide aged care?

Participant 16: Yeah, we do.

Interviewer: Where are they at with [the implementation]?

Participant 16: Ah, so we haven’t... We’ve only just really touched on the surface of um working with them, it’s quite tricky... You know, what’s a positive outcome in aged care? It’s quite different in terms of deterioration or whatever. So it ends up essentially coming back to more of quality, quality of service, but quality of life kind of measures. So that’s also, yeah, that was one it’s quite regulated and there’s a lot of reporting, quality reporting going on there. So, they are quite used to collecting data, but it’s a very different type I guess in terms of or very different concepts in that sense.

A lack of evaluation readiness in aged care and disability services was often linked back to challenges in establishing what should be measured and gaining service staff’s support for conducting these measurements. This issue frequently stemmed from a disconnect regarding the definition and measurement of “a positive outcome.” In aged care, there was a prevailing belief evaluation would not benefit the organisation, service, or recipient, as “deterioration” was seen as inevitable.

Participant 15: I remember when I had a conversation with our aged care manager at the time about outcomes measurement, she said. “But people die. And that’s a good outcome, right?” But it’s...How do they die? And how do we support them along that journey in terms of that transition? How do you make sure people are living with dignity?

This demonstrates the tension between wanting to deliver dignity in care, yet hesitating to conduct outcome measurement as demonstrating a positive impact was considered unlikely. Therefore, although inspection may uncover useful information, the null or negative finding is seen as too great a risk to the status quo. This contrasted with other services, particularly those for young people, where changing

recipient outcomes was seen as paramount, and therefore outcomes measurement and evaluation was very desirable. In these cases, the ability to demonstrate a positive impact of the service, and the potential to substantially improve service outcomes made evaluation a priority.

Participant 4: So yeah, so you have to make choices in terms of prioritising things that we believed had higher levels of contribution and innovation... You can't do everything, so we had simplistic kind of measures in aged care which mostly related to who we were working with and socioeconomic type things. But then in other things, we went deeper because there was a greater contribution and more to understand and learn and innovate and iterate.

For this participant and others, contribution and innovation were seen as factors which related to the ability of a service to create measurable change in recipient outcomes. Evaluation was often prioritised in such services including housing and homelessness, financial wellbeing, mental health, and drug and alcohol use. However, these areas were also considered easier to measure through the existence of associated quantitative indicators. On the other hand, the lack of existing quantitative measures to capture constructs related to aged care and disability services was often raised by participants.

Participant 17: We've got pretty well-established outcomes frameworks and I think housing and homelessness is probably further down that path and partly it's because it's just an easier thing to measure. Like is someone in sustained housing that suits them that they can afford in the longer term? You know that that's all quite measurable. ... What a good outcome looks like given the diverse range of disabilities that could be incorporated in disability services, is really difficult. Same with aged care. I think that's why you see auditing because we can talk about what standard services need to meet.

Although this frames outcomes in some areas as more "measurable", we suggest that the institutionalised environment surrounding these services has taken-for-granted beliefs and implied values that have driven a greater focus on measuring aspects we wish to change. Conversely, services not expected to change participant outcomes are not subject to the same expectations of pressures to conduct evaluations, making such effort unwarranted. However, as identified by this participant, inspection is still necessary, albeit directed towards the certification and auditing of service standards.

4.2.2 The Role and Shortcomings of Audit

Audit was routinely positioned as the appropriate inspection modality for programs where outcomes are not clearly defined. Therefore, audit and evaluation were often seen as incompatible. However, evaluators also framed audit around a rationale of assessing the more mundane, yet high risk aspects of the organisation. In comparing

evaluation and audit one senior evaluation manager described audit as looking “*for the extreme end of the spectrum, the thing that is maybe not that likely to occur, but if it were to occur, it sends the organisation into an absolute tailspin*” (Participant 2).

However, those directly involved in audit presented a broader picture of what audit offered their organisation. Rather than audits being used only to prevent risk they saw it as being important for inspecting day-to-day activities as well as protecting the dignity and wellbeing of service recipients. Indeed, the ability of an audit to occur at any time without warning is seen as a key benefit to surfacing how services are operating.

Participant 5: So we will typically know if we’re due for a full review audit the time period from which we can expect to receive that, but we don’t actually know until that day when that audit will occur, we’ll just arrive at work and a whole group of people will turn up on the doorstep and they’ll be ready to do what could be a four day review audit depending on the size of the site. So, yeah, and they come obviously with a particular agenda to review all standards that are relevant to whatever that area is.

Although the participant reflected that audit did safeguard against major issues, and that overall, being compliant meant that services should be of a higher standard, there were exceptions to this. They stated that a service could be fully compliant but not offer participants with choice or dignity, but also that a service could be non-compliant but provide the participant with the services they desire. This related back to the pre-definition of audit standards to meet upward accountability, rather than participant preferences.

Interviewer: Do you see that overall, audit, and having the standards well implemented, and having good results against those leads to increased quality of life outcomes?

Participant 5: I think that the answer is yes, based on my subjective assessment of what those standards cover. Whether that then is what those individuals feel is important or meaningful in their life I think is the question. Because effectively when we’ve got external auditors or ourselves trying to validate where people sit against those standards, what their experience is, what their perception is, umm, you know, we’re effectively asking them those questions within the parameters of demonstrating that we’re meeting those standards. And perhaps there’s a whole raft of things that are important to people in life that that possibly fall into another domain so.

The focus on service standards rather than recipient outcomes is a key historical feature of the social audit, however, recent disruptions in these sectors is leading to a change in the way audits are being conducted. Indeed, this new focus is driving “*significant change... massive regulatory change*” (Participant 6), directed towards measuring outcomes, effectiveness, and participant experience. Audit, although still

carrying an external and surprise element now encompasses processes more reminiscent of evaluation, in particular prioritising interviews with recipients.

Despite this overlap between the objectives of evaluation and audit, the two were not seen as complimentary or compatible. Indeed, despite the focus on outcomes, participants felt that evaluation could not be combined alongside audit because of the resource requirements in an already strained area. Rather, the “*system’s got to get to a point where it’s not seriously failing people before you can get to the point of optimising*” (participant 17).

4.2.3 Constitutive Effects of Service Recipients Characteristics on Inspection

To understand the systematic failures described by participants and the associated regulatory changes, it is necessary to provide additional context about the social service landscape in Australia. The system participants refer to is the quasi-market established around aged care, disability services, early childhood education, and training coordinated through marketised or “black-box” contracts (Considine 2022; Considine, Nguyen, and O’Sullivan 2018).

The majority of the 19 service areas listed in Figure 1 receive some government funding. In Australia, government funding can be broadly categorised into two contracting arrangements. The first, referred to as block-contracting, funding agreements, or grants, cover all reasonable and outlined expenses incurred during service delivery. This arrangement positions financial risk with the state, therefore not passing profit or loss to the service provider (Cunningham, Baines, and Charlesworth 2014; Malbon, Carey, and Dickinson 2018). Conversely, the service or market contract involves an agreed fixed unit price for service provision. Financial risk is allocated to the provider, creating opportunities for surplus through innovations that lower the actual cost of delivering a service unit (Considine 2022; Considine, Nguyen, and O’Sullivan 2018). However, this “black-box” funding arrangement is not always seen as benefiting service recipients with unethical and negligent practices observed in providers aspiring to maximise surplus, as uncovered by a “historic run of Royal Commissions” (Considine 2022, p. v). It is these commissions that have brought about the massive regulatory change particularly in disability, aged care, and early learning services. The lack of evaluation in these spaces was directly tied to the marketisation by one social sector leader.

Participant 4: So in other words, aged care, disability, even early learning, are operating, they’re all operating as markets, commercial competitive markets with for-profits and not-for-profits. So in that context, actually what’s more important is customer satisfaction, not outcomes, because of the nature of those market forces.

Several participants contrasted these market models with block funding services, especially those raised earlier as being the focus of evaluation such as family crisis, programs for teenagers and young people, mental health, and housing and homelessness services. Part of the rationale for why funders did not require evaluation in some areas was that innovation and improvement was drawn from the market mechanisms, rather than evidence and empiricism.

Interviewer: So do you think that's why it's not evaluated because it's, business as usual? They're not trying to test anything or... or it's not exciting?

Participant 2: Well, I think I don't think they're really trying to test anything or maybe they... I think they are trying to test things, but when they test them, they've kind of already decided what they wanna test because I think bottom line is they just wanna make it cheaper and cheaper and cheaper as time goes on.

This points to a fundamental difference in the framing of innovation or improvement between these service types; one aspiring towards effectiveness and the other towards cost-efficiency. However, when considering the differences in funding and inspection modalities, important insights were raised regarding the intended beneficiaries of these programs.

Participant 2: A lot of these sort of marginalised groups get an awful lot of funny project funding flung at them.

Interviewer: True. Yes.

Participant 2: And itsy bitsy teeny weeny project funding. Do you know what I mean? Ning nong funding. Sometimes it's the right funding because it's targeted funding and then that's great, that's exactly what we want. I think targeted is good. But sometimes it's quite piecemeal and bizarrely, in government, one of the things that I have noticed because [we] now provide really big government programs, believe it or not, it's the small projects that... by small I mean it might be \$600,000, is the thing that has the mandatory evaluation attached to it. Like, go figure?

Other participants who had worked in government raised this pattern as a source of irritation.

Participant 19: I mean, this was something that was a bug bear in treasury as well. Because you're often getting these new funding proposals and you're arguing over the few extra programs that are in or out [for evaluation] each year. But there's the whole base funding, which is like 80 percent, 90 percent of what you're funding for just the department to deliver its things that goes unreviewed and unmonitored quite regularly.

This “base funding” refers to large and enduring funding pools such as disability services, aged care, early childhood education, and welfare. A public sector

evaluation expert described these large pools of base funding which go unevaluated as “*perceived as a program that government has to deliver*” and “*anything that’s long term funded that doesn’t have a politically contentious item*” (participant 7).

From these descriptions, a dichotomy of social services emerges. First, long-standing, and non-contentious services which receive large yet marketised funding pools and are audited. Second, contentious programs with marginalised groups which receive small block funding and are evaluated. We examine this in the following discussion, illustrating how complex institutional environments lead to the adoption of different modalities of inspection.

5 Discussion

This research investigates how the characteristics of social services influence the modality and intensity of the inspection they undergo. Drawing on a framework of institutional theory (DiMaggio and Powell 1983; Greenwood 2008; Phillips et al. 2004; Zucker 1983), we propose that the adoption of inspection modalities is driven by taken-for-granted beliefs, schemas, and values in the environment rather than direct efficiency rationales.

Our findings show that within the Australian context the type of inspection depends on the social service model. Marketised models are more subject to compliance inspections and audits, while block grant models face effectiveness inspections through evaluation. This is a fascinating illustration of how institutional environments shape the blending of institutions as organisations strive to demonstrate legitimacy (Dahler-Larsen 2011; Meyer and Rowan 1977).

Through neo-liberal reforms, the primary question of inspection in the pursuit of accountability is whether the principal is getting value for money (Lapuente and Van de Walle 2020). For marketised models, as agents are allowed to innovate and reduce costs, the focus is on ensuring the service is delivered to specification (Considine 2001). As described by participant 4, this is the greatest challenge for marketised providers as they aim to respond to a “*government that is both fully deregulating and fully regulating at the same time*.” In these models, inspection assesses the agent (or service provider) itself, therefore blending the institution of inspection with the institution of capitalism with embedded managerial values.

Conversely, in small block grant funding, efficiency is achieved by providing more effective services that lower future public costs. Here, inspections focus on program outcomes, blending the institution of inspection with the institution of empiricism with embedded values of causality and attribution. This was illustrated powerfully in our qualitative interview data. For example, when contrasting the value of evaluation in early mental health intervention versus aged care, participant

12 observed a “*smaller relative change in impact for a younger person... will grow and you'll have a much bigger impact, whereas if you just help an older person, then your outcome isn't gonna last as long.*”

This contrast between the provision of services to recipients and the creation of change in recipients is demonstrated in the three areas with roughly equivalent application of compliance and effectiveness inspection: education and early learning, employment and training, and health services. Each of these areas encompassed two main types of programs: those utilised by a cross-section of society (e.g. childcare, tertiary training, and medical facilities), and those targeted as interventions for specific groups (e.g. inventions to address school drop-out, unemployment, and unhealthy behaviours). The former services are inspected via audit, while the latter interventions are inspected via evaluation.

5.1 Hypotheses and Future Research

Our data and prior scholarship lead us to propose the following hypotheses concerning the perceived deservedness of service recipients and the design and inspection modality of subsequent services. We recommend that these be examined in future research.

Building on the work of Meyer and Rowan (1977) and Dahler-Larsen (2011), and developing Considine's (2001, 2003, 2022) critical analyses of the Australian social services sector, we present the following hypothesis to account for this dynamic. First, social services that support individuals perceived as deserving are afforded

Table 1: Hypothesised dynamics between institutional status and inspection modalities.

Cross-sectional services (perceived as deserving)	Targeted Interventions (perceived as less deserving)
Afforded institutionalised status due to public perception of necessity.	Denied institutionalised status due to public perception of lower deservingness.
Face minimal scrutiny and inspection, as their necessity is rarely debated.	Subject to scrutiny and public debate, requiring inspection to bolster legitimacy.
Their continuity is assured, leading to stable inclusion in state budgets.	Their necessity is questioned, making ongoing funding and provision uncertain.
Marketisation is used to improve cost-efficiency in service delivery.	Marketisation is less feasible due to the contested nature of the services.
Certification and audit are the primary inspection modalities, aligning with market mechanisms.	Evaluation is the preferred inspection modality, as these services aim to generate measurable change in recipients' conditions.

institutionalised status. This status lowers the level of inspection they must engage with as there is little public debate regarding their necessity. These norms simultaneously increase the certainty of their ongoing provision and inclusion in state budgets, enabling the state to turn to marketisation to deliver the services with increased cost-efficiency. To inspect this marketisation, the market aligned inspection modalities of certification and audit are employed. Conversely, contentious services that support individuals perceived as less deserving are not afforded institutionalised status. There is scrutiny and public debate as to their necessity, and so inspection is adopted to bolster legitimacy. As the goal of these services is to generate a change in conditions and outcomes for the recipients (Parsell, Clarke, and Perales 2022), the empirically aligned inspection modality of evaluation is adopted. This categorisation is further illustrated in Table 1.

Therefore, the institutional environment shapes inspection modalities through embedded values and logics that differently frame which aspects of social services can and should be inspected: the provision of service itself, or the resulting changes for service recipients. Compliance and audit data is not less intensive than evaluation, rather it frames the object of inspection differently. This reflects the inherent epistemological divide between whose perspective on service performance is valued: upward accountability to the perspective of the purchaser, or downward accountability to the perspective of the recipient (Ahearn and Mai 2023; Ebrahim 2005).

This framing has caused further constitutive impacts in the Australian context. Namely, that the focus on upward rather than downward accountability in social services leads to mission drift and decreased service quality. Such decline has been highlighted by a raft of Royal Commissions (Considine 2022; Moulds 2021; Wade 2022). In Australia, Royal Commissions are established by the head of state to investigate matters of substantial public concern (Mintrom, O'Neill, and O'Connor 2021). These commissions, conducted on behalf of the Crown and often led by current or former members of the judiciary, can themselves be seen as a modality of inspection. Drawing on these associations with both the monarchy and the judiciary, they are granted paramount legitimacy (Dahler-Larsen 2011). Consequently, only organisations with severely threatened legitimacy in public discourse would be inspected through such a strongly institutionalised modality (Meyer and Rowan 1977).

The Royal Commissions into Violence, Abuse, Neglect and Exploitation of People with Disability, and Aged Care Quality and Safety, each emphasised the need for outcomes-focused care and participant inclusion to prevent the neglect and abuse they respectively identified in the marketised models (Considine 2022; Moulds 2021; Wade 2022). Interestingly, as expressed by the interview participants, these recommendations have led to changes in the conduct of audits to now focus on recipient perspectives through interviewing them. Such methodology is common in the evaluation space where lived experience and participatory approaches are

embedded by the constructionist paradigm (Vedung 2010). This increasing overlap between the methods of audit and evaluation will be a fascinating focus for future research. In particular, to what extent will the underlying values of capitalism and empiricism be maintained in the evolving application of inspection modalities.

The pattern linking recipient deservedness, funding confidence, marketisation levels, and legitimacy-based inspections may be applicable in other contexts. However, as our study is grounded in the Australian context, interpretations elsewhere should consider local values, cultures, and norms. Different value structures will result in varied manifestations of these patterns.

While previous research has investigated the motivations of not-for-profits to marketise their services (Suykens, De Rynck, and Verschueren 2019), our analysis points instead to the institutional factors behind the state's decision to marketise public services. The charities studied here are the most affluent in Australia, possessing considerably more capital to choose whether to participate in service provision. Participants discussed opting out of some forms of marketisation and public contracts deemed as not worthwhile for their organisation and communities. On the other hand, the not-for-profits in this sample providing aged care and disability services also represent the largest of these providers in Australia. It should be noted that these large not-for-profit providers were less criticised in the Royal Commissions for their service delivery than private and small providers. Nevertheless, the fact that these larger not-for-profit organisations also follow the market model and inspect via audit rather than evaluation indicates an institutional pattern that potentially accounts for issues among private providers.

The reliance on self-instigated reports of inspection published on organisation websites reflects a public threshold of accountability organisations seek to meet, rather than what may be selectively disclosed to funders (de Boer 2023). While this does not undermine the findings presented, understanding the difference between outward and internal accountability statements warrants further investigation. An embedded ethnographic approach would allow researchers to overcome the issue of self-disclosure, but also provide greater insight into the decision-making processes behind inspections including the adoption, deferral, or dismissal of inspection findings.

6 Conclusions

Our study provides a deeper understanding of how the institutional environment influences not only the inspection, but design, administration, and quality of services themselves. The institutional environment surrounding social services involves a complex interplay between the institutions of charity, the state, the market, and

empiricism. Each of these institutions brings with it taken-for-granted beliefs, schemas, and values. The values and afforded legitimacy embedded within services and stemming from the deservedness of recipients, whereby services are expected to care for, or change them, influences the modality of inspection adopted. Yet in a surprising consequence, recipients afforded the right to be cared for, are then faced with services inspected through modalities that historically did not consider their wellbeing and preferences, leading to suboptimal and even neglectful services. Therefore, current neo-liberal administration undermines the ability of social services to provide genuine care for recipients, deviating from the core values of the institution of charity. Through this insight we hope to prompt a critical re-examination of our expectations of social services, moving beyond the historical influence of deservedness.

Research funding: This work was supported by ARC Centre of Excellence for Children and Families over the Life Course (<http://dx.doi.org/10.13039/501100015792>, Project ID CE200100025).

References

- Ahearn, E.-R., and C. Mai. 2023. "The Nature of Measurement across the Hybridised Social Sector: A Systematic Review of Reviews." *Australian Journal of Public Administration* 1–20, <https://doi.org/10.1111/1467-8500.12616>.
- Ahearn, E.-R., and C. Parsell. 2024. "Imprinting and the Evolution of Evaluation: A Descriptive Account of Social Impact Evaluation Methodological Practice." *Evaluation* 30 (4): 568–88.
- Alkin, M. 2004. "Evaluation Roots." In *Evaluation Roots*, edited by M. Alkin, 374–80. Thousand Oaks, USA: SAGE Publications, Inc.
- Alkin, M., and S. M. Taut. 2002. "Unbundling Evaluation Use." *Studies in Educational Evaluation* 29 (1): 1–12.
- Armstrong, J. 2010. "Naturalistic Inquiry." In *Encyclopedia of Research Design*, 1st ed., 880–5. Thousand Oaks, California: SAGE.
- Australian Charities and Not-for-profits Commission. 2023. *Australian Charities Report*, 9th ed. <https://acnc.gov.au/tools/reports/australian-charities-report-9th-edition>.
- Bach-Mortensen, A. M., and P. Montgomery. 2018. "What Are the Barriers and Facilitators for Third Sector Organisations (Non-profits) to Evaluate Their Services? A Systematic Review." *Systematic Reviews* 7 (1). Article 1. <https://doi.org/10.1186/s13643-018-0681-1>.
- Barman, E. 2007. "What Is the Bottom Line for Nonprofit Organizations? A History of Measurement in the British Voluntary Sector." *Voluntas: International Journal of Voluntary and Nonprofit Organizations* 18 (2). Article 2. <https://doi.org/10.1007/s11266-007-9039-3>.
- Benjamin, L. M., A. Ebrahim, and M. K. Gugerty. 2023. "Nonprofit Organizations and the Evaluation of Social Impact: A Research Program to Advance Theory and Practice." *Nonprofit and Voluntary Sector Quarterly* 52 (1_suppl): 313S–352S.
- Bovens, M. 2007. "Analysing and Assessing Accountability: A Conceptual Framework." *European Law Journal* 13 (4): 447–68.

- Campbell, D. T. 1979. "Assessing the Impact of Planned Social Change." *Evaluation and Program Planning* 2 (1): 67–90.
- Christensen, T., and P. Lægreid. 2011. "Complexity and Hybrid Public Administration—Theoretical and Empirical Challenges." *Public Organization Review* 11 (4): 407–23.
- Clarke, A., and C. Parsell. 2022. "Resurgent Charity and the Neoliberalizing Social." *Economy and Society* 51 (2): 307–29.
- Considine, M. 2001. *Enterprising States: The Public Management of Welfare-to-Work*. Cambridge: Cambridge University Press.
- Considine, M. 2003. "Governance and Competition: The Role of Non-profit Organisations in the Delivery of Public Services." *Australian Journal of Political Science* 38 (1): 63–77.
- Considine, M. 2022. *The Careless State: Reforming Australia's Social Services*. Melbourne, Australia: Melbourne University Press.
- Considine, M., P. Nguyen, and S. O'Sullivan. 2018. "New Public Management and the Rule of Economic Incentives: Australian Welfare-to-Work from Job Market Signalling Perspective." *Public Management Review* 20 (8): 1186–204.
- Cooper, D. J., and M. J. Sherer. 1984. "The Value of Corporate Accounting Reports: Arguments for a Political Economy of Accounting." *Accounting, Organizations and Society* 9 (3–4): 207–32.
- Cordery, C. J., and R. Sinclair. 2013. "Measuring Performance in the Third Sector." *Qualitative Research in Accounting and Management* 10 (3/4): 196–212.
- Cunningham, I., D. Baines, and S. Charlesworth. 2014. "Government Funding, Employment Conditions, and Work Organization in Non-profit Community Services: A Comparative Study." *Public Administration* 92 (3): 582–98.
- Dahler-Larsen, P. 2011. *The Evaluation Society*. Stanford, California: Stanford University Press.
- de Boer, T. 2023. "Updating Public Accountability: A Conceptual Framework of Voluntary Accountability." *Public Management Review* 25 (6): 1128–51.
- De Waele, L., T. Polzer, W. A. van, and L. Berghman. 2021. "“A Little Bit of Everything?” Conceptualising Performance Measurement in Hybrid Public Sector Organisations through a Literature Review." *Journal of Public Budgeting, Accounting and Financial Management* 33 (3): 343–63.
- Deegan, C. 2002. "Introduction: The Legitimising Effect of Social and Environmental Disclosures – A Theoretical Foundation." *Accounting, Auditing & Accountability Journal* 15 (3): 282–311.
- Dees, J. G. 2012. "A Tale of Two Cultures: Charity, Problem Solving, and the Future of Social Entrepreneurship." *Journal of Business Ethics* 111 (3): 321–34.
- Dhanani, A., and C. Connolly. 2012. "Discharging Not-for-profit Accountability: UK Charities and Public Discourse." *Accounting, Auditing & Accountability Journal* 25 (7): 1140–69.
- DiMaggio, P. J., and W. W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48 (2): 147–60.
- Ebrahim, A. 2005. "Accountability Myopia: Losing Sight of Organizational Learning." *Nonprofit and Voluntary Sector Quarterly* 34 (1): 56–87.
- Ebrahim, A. 2019. *Measuring social Change: Performance and Accountability in a Complex World*. Stanford: Stanford University Press.
- Eikenberry, A. M., and J. D. Kluver. 2004. "The Marketization of the Nonprofit Sector: Civil Society at Risk?" *Public Administration Review* 64 (2): 132–40.
- Epstein, M., E. Flamholtz, and J. J. McDonough. 1976. "Corporate Social Accounting in the United States of America: State of the Art and Future Prospects." *Accounting, Organizations and Society* 1 (1): 23–42.
- Fairclough, N. 1992. *Discourse and Social Change*. Cambridge, England: Polity Press.

- Gibbon, J., and C. Dey. 2011. "Developments in Social Impact Measurement in the Third Sector: Scaling up or Dumbing Down?" *Social and Environmental Accountability Journal* 31 (1). Article 1. <https://doi.org/10.1080/0969160X.2011.556399>.
- Gray, R. 2001. "Thirty Years of Social Accounting, Reporting and Auditing: What (If Anything) Have We Learnt?" *Business Ethics: A European Review* 10 (1). Article 1. <https://doi.org/10.1111/1467-8608.00207>.
- Gray, R., C. Dey, D. Owen, R. Evans, and S. Zadek. 1997. "Struggling with the Praxis of Social Accounting: Stakeholders, Accountability, Audits and Procedures." *Accounting, Auditing & Accountability Journal* 10 (3). Article 3. <https://doi.org/10.1108/09513579710178106>.
- Gray, R., D. Owen, and C. Adams. 2009. "Some Theories for Social Accounting? A Review Essay and a Tentative Pedagogic Categorisation of Theorisations around Social Accounting." In *Sustainability, Environmental Performance and Disclosures*, Vol. 4, edited by M. Freedman, and B. Jaggi, 1–54. Leeds: Emerald Group Publishing Limited.
- Greenwood, R. 2008. *The SAGE Handbook of Organizational Institutionalism*. Los Angeles: SAGE.
- Greenwood, R., M. Raynard, F. Kodeih, E. R. Micelotta, and M. Lounsbury. 2011. "Institutional Complexity and Organizational Responses." *The Academy of Management Annals* 5 (1): 317–71.
- Hillman, A., W. Tadd, S. Calnan, M. Calnan, A. Bayer, and S. Read. 2013. "Risk, Governance and the Experience of Care." *Sociology of Health & Illness* 35 (6): 939–55.
- Holstein, J. A., and J. F. Gubrium. 1995. *The Active Interview*, Vol. 37. Thousand Oaks, California: Sage.
- Hsieh, H.-F., and S. E. Shannon. 2005. "Three Approaches to Qualitative Content Analysis." *Qualitative Health Research* 15 (9): 1277–88.
- Johnsson, M. C., M. Pepper, O. M. Price, and L. P. Richardson. 2021. "“Measuring Up”: A Systematic Literature Review of Performance Measurement in Australia and New Zealand Local Government." *Qualitative Research in Accounting and Management* 18 (2): 195–227.
- Katz, M. B. 1986. *In the Shadow of the Poorhouse: A Social History of Welfare in America*. New York: Basic Books.
- Kendall, J., and M. Knapp. 2000. "Measuring the Performance of Voluntary Organizations." *Public Management: International Journal of Reality Therapy* 2 (1): 105–32.
- Kurzban, R., M. N. Burton-Chellew, and S. A. West 2015. The Evolution of Altruism in Humans. *Annual Review of Psychology* 66: 575–99.
- Lakoff, G. 2014. "The All New Don't Think of an Elephant." In *Know Your Values and Frame the Debate*, Revised and updated edition. White River Junction, Vermont: Chelsea Green Publishing.
- Lapuente, V., and S. Van de Walle. 2020. "The Effects of New Public Management on the Quality of Public Services." *Governance* 33 (3): 461–75.
- Lingane, A., and S. Olsen. 2004. "Guidelines for Social Return on Investment." *California Management Review* 46 (3): 116–35.
- Malbon, E., G. Carey, and H. Dickinson 2018. Accountability in Public Service Quasi-markets: The Case of the Australian National Disability Insurance Scheme. *Australian Journal of Public Administration* 77(3): 468–81.
- Meyer, J., and B. Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83 (2). <https://doi.org/10.1086/226550>.
- Mintrum, M., D. O'Neill, and R. O'Connor. 2021. "Royal Commissions and Policy Influence." *Australian Journal of Public Administration* 80 (1): 80–96.
- Moulds, S. 2021. "Royal Commission into Aged Care Quality and Safety: Time for a Paradigm Shift to Protect the Human Rights of Older Australians." *Bulletin (Law Society of South Australia)* 43 (6): 14–7.
- Moxham, C. 2014. "Understanding Third Sector Performance Measurement System Design: A Literature Review." *International Journal of Productivity and Performance Management* 63 (6): 704–26.

- O'Brien, A. 2015. *Philanthropy and Settler Colonialism Anne O'Brien*. Basingstoke, Hampshire: Palgrave Macmillan.
- Onyx, J., L. Cham, and B. Dalton. 2016. "Current Trends in Australian Nonprofit Policy." *Nonprofit Policy Forum* 7 (2): 171–88.
- Parsell, C., A. Clarke, and F. Perales. 2022. *Charity and Poverty in Advanced Welfare States*. Oxon: Routledge.
- Patton, M. Q. 2008. *Utilization-Focused Evaluation*, 4th ed. Thousand Oaks: Sage Publications.
- Phillips, N., T. B. Lawrence, and C. Hardy. 2004. "Discourse and Institutions." *Academy of Management Review* 29 (4): 635–52.
- Powell, M. 2007. *Understanding the Mixed Economy of Welfare*. Bristol, UK: Policy Press.
- Quadagno, J. 1987. Theories of the Welfare State. *Annual Review of Sociology* 13: 109–28.
- Ramanathan, K. V. 1976. "Toward a Theory of Corporate Social Accounting." *The Accounting Review* 51 (3). Article 3.
- Reinertsen, H., K. Bjørkdahl, and D. McNeill. 2022. "Accountability versus Learning in Aid Evaluation: A Practice-Oriented Exploration of Persistent Dilemmas." *Evaluation* 28 (3): 356–78.
- Salamon, L. M. 1987. "Of Market Failure, Voluntary Failure, and Third-Party Government: Toward a Theory of Government-Nonprofit Relations in the Modern Welfare State." *Journal of Voluntary Action Research* 16 (1–2): 29–49.
- Sawhill, J. C., and D. Williamson. 2001. "Mission Impossible? Measuring Success in Nonprofit Organizations." *Nonprofit Management and Leadership* 11 (3): 371–86.
- Spence, C. 2009. "Social Accounting's Emancipatory Potential: A Gramscian Critique." *Critical Perspectives on Accounting* 20 (2). Article 2. <https://doi.org/10.1016/j.cpa.2007.06.003>.
- Stake, R. E. 2001. "A Problematic Heading." *American Journal of Evaluation* 22 (3): 349–54.
- Suykens, B., F. De Rynck, and B. Verschueren. 2019. "Examining the Influence of Organizational Characteristics on Nonprofit Commercialization." *Nonprofit Management and Leadership* 30 (2): 339–51.
- Tan, H. T. R., and G. Harvey. 2016. "Unpacking the Black Box: A Realist Evaluation of Performance Management for Social Services." *Public Management Review* 18 (10): 1456–78.
- Vaismoradi, M., J. Jones, H. Turunen, and S. Snelgrove. 2016. "Theme Development in Qualitative Content Analysis and Thematic Analysis." *Journal of Nursing Education and Practice* 6 (5): p100.
- Vedung, E. 2010. "Four Waves of Evaluation Diffusion." *Evaluation* 16 (3): 263–77.
- Wade, N. 2022. "The Disability Royal Commission: An Instrument for Law Reform and Public Policy Change." *Bulletin (Law Society of South Australia)* 44 (10): 6–7.
- Weiss, C. H. 1999. "The Interface between Evaluation and Public Policy." *Evaluation* 5 (4). Article 4. <https://doi.org/10.1177/135638909900500408>.
- Williams, P. F. 1987. "The Legitimate Concern with Fairness." *Accounting, Organizations and Society* 12 (2): 169–89.
- Young, D. R. 2002. "The Influence of Business on Nonprofit Organizations and the Complexity of Nonprofit Accountability: Looking inside as Well as outside." *The American Review of Public Administration* 32 (1): 3–19.
- Zucker, L. G. 1983. "Organizations as Institutions." *Research in the Sociology of Organizations* 2: 1–47.