

RESEARCH ARTICLE

WILEY

Balancing perspectives on performance: “Measurement from the inside” and “measurement from the outside”

Boram Lee¹  | Rodolfo Sejas-Portillo² | Ian Fraser³

¹UniSA Business, University of South Australia, Adelaide, South Australia, Australia

²Economics Division, Stirling Management School, University of Stirling, Stirling, UK

³Accounting and Finance Division, Stirling Management School, University of Stirling, Stirling, UK

Correspondence

Boram Lee, UniSA Business, University of South Australia, Adelaide, SA 5000, Australia.

Email: boram.lee@unisa.edu.au

Abstract

This article investigates the nature of tensions “on the ground” between the internal and external stakeholders of arts organizations in terms of performance measurement. Based on the qualitative analysis of 19 interviews, the performance measurement practices of two different-sized arts organizations highlight internal and external stakeholders’ contrasting perspectives on a number of measurement dimensions. In endeavoring to understand tensions between internal and external stakeholders, the article highlights the main differences which result from seemingly opposed ways of knowing. Internal stakeholders tend more to reflect “*phronesis*,” based on value-rationality, while external stakeholders are more inclined to technical (“*techne*”), or analytical (“*episteme*”) knowledge based on instrumental rationality. Nevertheless, there is some evidence of positive engagement between internal and external stakeholders. The article argues that in order to mitigate tensions, internal and external stakeholders should aim for culturally embedded understanding through

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Nonprofit Management & Leadership* published by Wiley Periodicals LLC.

evaluation-focused dialogue. In providing both theoretical contributions to nonprofit literature and managerial implications, the article offers an original perspective that challenges existing practice and calls for greater understanding of conflicting interests when different stakeholders are involved in performance measurement.

KEYWORDS

arts organizations, funding agencies, performance measurement, tensions

1 | INTRODUCTION

This article investigates the nature of the tensions between internal and external stakeholders of arts organizations which result from conflicting perspectives on how to measure performance. Previous nonprofit performance measurement studies find that tensions exist between theory and practice (Schuster, 1996); between measures and strategy (Johnston & Pongatchat, 2008); between historic and more recent practice (Benjamin, 2010), all due principally to the challenges in applying quantitative indicators in response to funders' demands for accountability. Although many studies highlight the role of funders in driving performance measurement by nonprofits and the tensions between internal and external stakeholders (Benjamin, 2010; Carman, 2007, 2009; Moxham, 2010), few studies explore what drives such tensions. Furthermore, although performance measurement challenges have been well captured for nonprofits generally (Carman, 2007, 2009; Carman & Fredericks, 2010; Cavalluzzo & Ittner, 2004), this does not apply to arts organizations.

Very few studies investigate how performance measurement operates in arts management practice, for whom and by whom it is done, and how it is perceived by different stakeholders (Gilhespy, 1999; Turbide & Laurin, 2009). Despite the important role of funders highlighted in the literature, no study articulates nuanced arguments for why and how arts organizations' internal and external stakeholders perceive performance measurement practices differently and the causes of the associated tensions. Further, no study has investigated ways of resolving such tensions. Wider scholarship on nonprofits highlights tensions between different stakeholders due to conflicting expectations and recommends that managers should engage in evaluation-based dialogue with external stakeholders (Becker et al., 2011; Herman & Renz, 1997, 1998). Nevertheless, the tensions between internal and external stakeholders "on the ground" for arts organizations have not been fully investigated. We respond by exploring the following question: *How do tensions between the internal and external stakeholders of arts organizations reflect conflicting perspectives on performance measurement?*

Informed by the two alternative perspectives on inquiry in the organizational sciences proposed by Evered and Louis (1981, p. 385), we operationalize these perspectives as "measurement from the inside" and "measurement from the outside." In so doing, we are cognizant of two seemingly opposed "ways of knowing," discussed by Flyvbjerg (2001); "*phronesis*"

underpinning the one and “*episteme*” and “*techne*” the other. “*Phronesis*” goes beyond technical (“*techne*”) or analytical (“*episteme*”) knowledge and Flyvbjerg argues that its greater importance is because of it being “that activity by which instrumental rationality is balanced by value-rationality” (Flyvbjerg, 2001, p. 4). Polkinghorne (2004) draws on these concepts in discussing the practice of care, arguing that “everyday practices do not ordinarily issue from conscious, rational calculation; instead, they flow from background understandings that are culturally embedded” (Polkinghorne, 2004, p. 152).

Although not Polkinghorne’s concern, this would appear to be a particularly apt description of the way arts organizations (should) operate and Polkinghorne advocates “*phronesis*” as an approach to decision-making in “human” practices which involve multiple values, goals, interests, and emotionality. We suggest a priori that “insiders” are more likely to reflect elements of “*phronesis*” in their attitudes to performance measurement than are “outsiders” who are more likely to exhibit “*episteme*” or “*techne*.” Our ultimate purpose in contrasting the two perspectives is to mitigate tensions between internal and external stakeholders through understanding these epistemological paradigms. This may help legitimate arts organizations’ search for more appropriate measures based on artistic understanding (“*phronesis*”) while also taking into account external stakeholders’ requirements for technical (“*techne*”) or analytical (“*episteme*”) “correctness” (Evered & Louis, 1981).

Arts organizations often struggle with external pressures due to limited capacity and performance measurement resources (Abdullah et al., 2018). We account for different contexts to gain a view of the different practices between arts organizations with varying levels of capacity and resources and explore how these factors might contribute to the extent to which they engage in performance measurement practices (Carman, 2009; Labaronne, 2017). Evered and Louis (1981, p. 390) define “context” as the complex fabric of local culture, people, resources, purposes, earlier events, and future expectations that constitute the time-and-space background of a specific situation.

We find that internal and external stakeholders’ perspectives differ on multiple dimensions of measurement (Evered & Louis, 1981) leading to tensions and in them employing opposed “ways of knowing” through the lens of practice theory (Flyvbjerg, 2001; Polkinghorne, 2004). Our research makes important contributions highlighting both similarities and differences in the particular tensions experienced by arts organizations as compared to other nonprofits. Although tensions between internal and external stakeholders due to institutionalized performance measurement systems are as evidenced in other nonprofits, the subjective and intangible character of their long-term artistic objectives and achievements lead to particular tensions being experienced within arts organizations. Further, unlike many other nonprofits, arts organizations are either micro or small-sized and heavily dependent on government funding; thus, they tend to lack the capacity and resources to invest in appropriate and effective performance measurement systems. Our study has implications for those who manage arts organizations in terms of mitigating tensions between internal and external stakeholders. Although our findings are grounded in the arts and cultural sector, our study offers insights for other nonprofits faced with tensions between multiple stakeholders with conflicting objectives while managing multiple revenue streams. A deteriorating environment of declining government subsidies identifies our study as being relevant to the nonprofit sector generally.

The following sections, in turn, review the relevant literature, explain the research methods used, report our findings, discuss, and interpret these findings and, finally, present our conclusions.

2 | PERFORMANCE MEASUREMENT FOR NONPROFITS

Performance measurement by nonprofits is a response to increased demands for accountability, transparency, and financial responsibility (Benjamin, 2010; Carman, 2009; Carman et al., 2008; Lacey et al., 2012). Transition from the one-dimensional approach of managing by numbers to an emphasis on both financial and non-financial indicators using the Balanced Scorecard (Kaplan & Norton, 1996) has been an important development. Accounting for multiple dimensions of performance is reflective of the multiple stakeholders who have an interest in various aspects of nonprofits' mission, strategy, and performance (Atkinson et al., 1997; Kaplan, 2001). Previous studies, however, highlight the difficulties and methodological challenges involved in measuring nonprofits' performance especially when guiding strategy and capturing mission accomplishment (Becker et al., 2011). Highlighting the weak link between performance measurement and performance improvement, Moxham (2010) identifies resource-intensive measurement practices, funder-focused measurement criteria, confusing and inconsistent use of terminology, inability to plan for the medium- or long-term due to insecure funding, and internal resistance as factors detracting from the usefulness of performance measurement by nonprofits.

While there is a need for nonprofits to measure performance in terms of their "mission," the foundation of their organizational ethos, rather than financial outcomes (Cochran et al., 2008; Pandey et al., 2017), Sawhill and Williamson (2001, p. 103) find that "very few nonprofits have systematically linked their metrics to their mission." The alignment of performance measurement with strategic goals is likely to be effective; however, only when relevant measurement criteria are developed from organizational strategy (Kaplan & Norton, 1996). Johnston and Pongatichat (2008) argue that in order to mitigate tensions between performance measures and strategy, nonprofits undertake three coping strategies: "do-nothing," "pseudo-realigning," and "distracting." Frumkin (2002) criticizes those nonprofits which tend to rely on informal, nongeneralizable, and hard-to-use reports, lacking in rigor and professionalism, believing that "the important" is not easy to measure. Eikenberry and Kluver (2004) caution against the "marketization" of nonprofits' actions, structures, and philosophies with an emphasis on effectiveness, accountability and legitimacy which intersect and interact. To address discrepancies between mission and performance measures, tailoring more nuanced measures to individual nonprofits' missions, goals and social contexts is fundamental to developing effective appraisal (Levy & Williams, 2004; Sawhill & Williamson, 2001).

Nonprofit measurement is complicated due to multiple stakeholders demands, goals and objectives leading to conflicts over what constitutes success and how this should be measured (Becker et al., 2011; Herman & Renz, 1997). Herman and Renz (1999) argue that socially constructed nonprofit performance is comparative and multidimensional, responsive to different stakeholder expectations, and rarely employs universally applicable or unanimously approved performance measures. Herman and Renz (1998) find that more effective nonprofits have boards with greater social prestige and feature professional procedures and strategies. Emphasizing the value of developing evaluation-focused dialogue with stakeholders, Herman and Renz (1997, 1998) recommend that nonprofit managers engage in ongoing dialogue focused on the assessment criteria used by different stakeholders to inform and shape their expectations. Herman and Renz (2004) claim that constructive dialogue may enable nonprofit managers to distinguish the "right fit" for each stakeholder request rather than pursuing a more inflexible emphasis on doing things the "right way."

The limited capacity and resources which nonprofits are able to allocate to evaluation in response to demands for performance measurement by funding bodies has been noted as a persistent issue. Using cluster analysis, Carman and Fredericks (2010) identify three types of nonprofits: those which are “satisfied with their evaluation efforts” while having limited time to devote to evaluation; those which “struggle despite internal support” from management, board and staff, but which have some capacity to implement evaluation systems; and those which “struggle across the board with implementation challenges” due to insufficient staff, funding, time or expertise. Carman and Frederick highlight variability in the evaluation experiences of nonprofits, their need to enhance capacity and the importance of information technology. According to Cavalluzzo and Ittner (2004), the lack of technical capability to generate timely and relevant information, weak management commitment and a lack of employee training leads to measurement failure especially of qualitative outcomes, arguably those which are most relevant. This results in excessive emphasis on financial performance indicators. In reviewing US federal legislation and mandates calling for effective philanthropy and improvements in nonprofits’ evaluation practices, Carman et al. (2008) recommend accessible training to enhance evaluation capacity. Carman (2009) argues that nonprofits should use their limited resources more strategically; while most funders require evaluation and performance measurement data, not all engage in extensive monitoring or require detailed reporting and evaluation. Often nonprofits’ peer-reputations influences funders’ decisions (Carman, 2007).

3 | PERFORMANCE MEASUREMENT FOR ARTS ORGANIZATIONS

Subsidized arts organizations, both at local and national levels, are regulated by policies emphasizing “hard” evidence and instrumentalism (Rentschler, Lee, & Subramaniam, 2021). The centrality of accounting practices to government funding decisions (Donovan & O’Brien, 2016; Rentschler et al., 2022; Rentschler, Lee, & Fillis, 2021) has escalated tensions between “creativity” within the arts and cultural sector and “control” by government. Chiaravalloti (2014) finds that the arts and cultural sector has engaged more with financial accounting, focusing principally on producing financial information for external stakeholders rather than management accounting which is concerned with producing both financial and non-financial information for internal stakeholders in order to assist decision-making, allocation of resources, monitoring, control and reward of performance (Atkinson et al., 1997). Previous studies on performance measurement for arts organizations stress underlying tensions between competing, or even mutually exclusive values, some of which may be exclusively artistic or market-oriented (Labaronne & Piber, 2020; Turbide & Laurin, 2009; Z. G. Voss & Voss, 2000). The adoption of rationalist business approaches often leads to conflict at both organizational and personal levels (Krug & Weinberg, 2004).

Although Chiaravalloti and Piber (2011) argue that the evaluation of artistic outcomes and impacts should form an integral part of performance measurement, translating qualitative objectives into valid and reliable performance indicators or providing robust accounting evidence of what is delivered to individuals and society is challenging (Boorsma & Chiaravalloti, 2010). Summarizing the multiple objectives of arts organizations into a relatively small number of indicators privileges financial results or audience numbers, rather than artistic achievement due to the challenges of measuring qualitative outcomes (Gilhespy, 1999; Schuster, 1996; Turbide & Laurin, 2009).

A recent attempt to measure artistic quality is *Culture Counts*, a metrics-based assessment system that collates data on the different qualities of cultural events based on audience surveys (Gilmore et al., 2017), but it has engendered much debate. While Throsby (2017) views *Culture Counts* as a response to demands for assessment of artistic quality and as a legitimate exercise trying to make sense of people's judgments after experiencing cultural events, Phiddian et al. (2017) criticize it for inviting political manipulation and for encouraging a panoptic analytical view devoid of meaningful judgment. Labaronne and Piber (2020) reject reliance on external measures of artistic quality such as decontextualized audience surveys and positivist economic analyses which fail to represent artistic effort adequately; these are contextual, subjective, intangible, long-term and often unquantifiable. In response, Labaronne and Tröndle (2021) suggest a framework reflecting the real-life complexities of intertwined management and artistic practice. Boerner (2004), in attempting to measure the artistic quality of opera companies, identifies various criteria. Nørreklit (2011), however, commends the artistic director of the Royal Danish Opera who facetiously suggests using a hydrometer in order to measure the strength of audiences' weeping, thereby ridiculing attempts to measure such symbols of quality.

While many arts organizations produce metric-saturated performance reports, Meyrick et al. (2018) question their usefulness. Although the benefits of cultural activities are often hard to articulate in quantitative, especially monetized, terms, assessments are skewed toward ease of use, and when applied uncritically, may mislead (Meyrick et al., 2018). Proposing an artistic-mission-led evaluation model, Boorsma and Chiaravalloti (2010) argue that the Balanced Scorecard has been imposed on the arts and cultural sector without fully acknowledging its distinctive nature. For instance, the measurement metrics used may fail to recognize the dimensions recognized by multiple stakeholders which may have conflicting objectives (Balsler & McClusky, 2005; G. B. Voss et al., 2000).

According to Turbide and Laurin (2009), funding agencies are regarded as the most important stakeholder by arts organizations to whom they are accountable; this affects their approach to performance measurement and their choice of performance indicators. Peterson (1986, p. 256) argues that, compared to other patrons, government funding agencies require far more "formalized" and "standardized" documentation and evaluation, affecting both arts organizations' access to government funds and their organizational structure (Froelich, 1999; Schuster, 1996). Similarly, Schuster (1996) draws attention to performance indicators which can serve not only evaluative, but also "attention-directing," purposes. He argues that quantitative indicators may affect the evaluative and monitoring behavior of arts organizations adversely. Over-emphasis on performance indicators may eventually lead to a loss of *raison d'être* and crowd-out artistic practice (Weinstein & Bukovinsky, 2009). Z. G. Voss and Voss (2000) and Evans (2000) find that there is a negative impact on the financial performance of arts organizations when they become too customer focused.

4 | METHODS

4.1 | Study context

We adopt a "real-life" and holistic perspective, and thus we favor a qualitative research method with a realist orientation (Miles et al., 2018). The qualitative method was deemed the most appropriate due to the exploratory and descriptive nature of our research and its contemporary focus. We conduct a multiple-case study, and our sample of arts organizations is drawn from a single small city in Scotland (population less than 100,000).

As we are interested in the tensions between internal and external stakeholders, we select only arts organizations which receive funding from external stakeholders including local councils and national government funding agencies. A purposeful sampling strategy was adopted aiming for high variation (Patton, 2014) accounting for different contexts of the sampled cases, with the legal form, size (i.e., resource mobilization volumes including funding), capacity, and time-in-business (i.e., age) being our main variation conditionals (Carman, 2009). Organizations with different legal forms operate under different performance measurement and reporting regulatory requirements. Organizations of different sizes tend to use different performance indicators and their time-in-business normally determines the degree of professionalization in their performance measurement and reporting procedures and processes (Garengo et al., 2005).

Two arts organizations which meet these criteria were selected to gain an extended view of the prevalence and importance of different practices within the sector (Eisenhardt & Graebner, 2007). The pseudonyms, *Stage* and *Connect*, are used for these organizations in order to ensure the anonymity of participants and, similarly, the city's name is not disclosed. *Stage* is a well-established organization with multiple income streams including government funding and significant private donations while *Connect* is a small, recently formed organization with limited access to funding.

As presented in Table 1, *Stage*, established in 1971 as a multi-arts venue, offers various programs including live performances, film screenings, art exhibitions, and workshops, attracting over 160,000 visitors each year and with a turnover of £1.68 m according to its annual report for the year ended 31 March Throsby, 2017. The mission of *Stage* is “to be a thriving arts center for everyone.” *Stage* also produces and co-produces new performances, festivals, and events, some

TABLE 1 Backgrounds of arts organizations.

	Stage	Connect
Year established	1971	2012
Number of employees	33	8
Mission	“To be a thriving arts center for everyone”	“To make the city a lively and culturally diverse place to live, work and visit”
Key stakeholders	<ul style="list-style-type: none"> • Artists • Critics/peers • Community • Industry partners (university, sponsors, etc.) • Local council • National arts funding body 	
Key resources	<ul style="list-style-type: none"> • Public (grants, funding, etc.) • Market (ticket sales, commission, etc.) • Private (donation, volunteering, etc.) 	
Current legal form	Charity (independent limited company)	
Years in legal form	Since 2008 (previously a university department)	Since 2016 (previously a community interest company)

of which tour the UK and abroad. It currently has 7 board members and 33 full and part-time staff. It initially operated as a University Department and became an independent registered charity in 2008. *Stage* is one of the Regularly Funded Organizations (RFO) of the national funding agency and its premises are subsidized by the university in which it is located.

Connect was established in 2012 as a Community Interest Company (CIC) but became a registered charity in 2016. The mission of *Connect* is “to make the city a lively and culturally diverse place to live, work and visit.” Its activities include art exhibitions, outdoor music events, film screenings, and workshops in different venues around the city with the aim of bringing cultural and learning opportunities to local communities. *Connect* also manages a retail shop which sells works produced by local artists. According to its annual report for the year ended 30 June 2017, its turnover was £41,516. *Connect* currently has 8 staff employed both full and part-time and over 25 volunteers including a “shadow” board. *Connect* is supported by both the national funding agency via its “Targeted Funding” initiative, and by its local authority which grants the use of building premises at subsidized rates. These two organizations are different in terms of their activities, size, establishment, and reputation, but similar in terms of being subject to the same local and national government policies. National government funding for arts organizations in Scotland is centralized through a single agency, *Creative Scotland*, which awards funds primarily from the UK National Lottery, and local councils provide additional financial support for arts organizations.

4.2 | Data collection and analysis

We follow the theoretical framework of Evered and Louis (1981) in order to capture “inside” and “outside” perspectives within arts organizations with the practice theory of Flyvbjerg (2001) and Polkinghorne (2004) underpinning our data collection and analysis. The research design provided clarity and focus during data collection and analysis. Nonetheless, the data collection procedures were receptive to propositions which might be in conflict or missing from the theoretical discussion. Similarly, during the data analysis caution was applied to avoid interpreting data out of context when comparing the cases in the context of our theoretical discussion. This helped to avoid different interpretations and non-comparable data across the cases (Miles et al., 2018; Yin, 2015).

Semi-structured qualitative interviews with the stakeholders of each organization were used as the collection method for primary data. In order to identify potential interviewees; internal stakeholders' (directors, managers, board members, and members of staff) roles were considered based on: (a) their strategic management activities; and (b) the operational activities for which performance measurement practices take place. Similarly, for external stakeholders, the context of the arts and culture sector in Scotland is considered and industry partners and funders from local councils and national funding agencies are identified as potential interviewees. The purpose of the interviews was to gather detailed information about perspectives on performance measurement practices for the selected arts organizations. Importantly, we do not consider audiences, volunteers, or other community members for interviews since they do not impact the performance measurement or reporting activities of organizations directly. We focus only on “practitioners” (Carman, 2007; Herman & Renz, 1998) who make judgments about the effectiveness of performance measurement criteria.

As shown in Table 2, 19 individuals were interviewed (12 internal and 7 external stakeholders, including 2 industry partners and 5 representatives from local and national funding

TABLE 2 Interviewees' background and length of interviews.

Interviewees	Stage	Min	Connect	Min
Internal stakeholder	The chair of the board	56	Trustee	36
	Artistic director	46	Director	45
	Operations manager	54	Creative producer	45
	Financial manager	49	Project manager	45
	Venue project manager	37	Event manager	51
	Admin & development manager	37	Head of marketing	31
External stakeholder	Local council staff	53	Local council officer	26
	Corporate partner	20	Local council director	60
	National funding agency director	58	National funding agency staff	35
			Industry partner	27

agencies), with each interview taking about 45 minutes. Each interviewee was asked about their organizational role, and how that was encapsulated within performance measurement practice.

In addition, 20 documents were collected (e.g., Annual Reports, Partnership Plan, Year End Briefing, and End of Project Monitoring Report) relating to the period 2013 to 2018, with reference to the performance measurement practices of each organization and that of funders, in order to support, confirm and complement the information acquired through the interviews.

Digitally recorded interviews were transcribed and then, together with the documents collected, they were coded, classified, and analyzed using NVivo. Due to the exploratory and descriptive nature of the study, the “pattern matching” method of content analysis was employed (Patton, 2014). Our analysis is informed by the two alternative perspectives on inquiry in the organizational sciences proposed by Evered and Louis (1981, p. 385). We operationalize these perspectives, for our purposes, as “measurement from the inside” and “measurement from the outside.” Data was first coded to capture the recurrent and distinctive themes in examining perspectives of internal and external stakeholders, highlighting similarities and differences in order to capture tensions, and descriptions were then created from these themes. Our analysis is also guided by the practice theory of Flyvbjerg (2001) which describes two opposed “ways of knowing” by means of the concepts of “*phronesis*,” “*episteme*” and “*techne*”: highlighting the value-rationality of culturally embedded understandings (“*phronesis*”) and conscious, rational calculation based instrumental rationality (“*episteme*” and “*techne*”) (Polkinghorne, 2004). We expect a priori that “inside” stakeholders exhibit *phronesis* in their approach to performance measurement while those on the “outside” are more likely to display *techne* or *episteme*. Nevertheless, we anticipated, in accordance with Labaronne (2017), some measurement of positive engagement between internal and external stakeholders.

The emerging patterns and themes were then identified and interpreted qualitatively. The patterns and themes identified were triangulated from different information sources (e.g., performance reports, policy reports, annual reports) and, within each source, from different interviewees (Patton, 2014). Although coding was undertaken individually by the researchers, differences in opinion were reconciled.

5 | FINDINGS

5.1 | Measurement from the inside

Directors of both organizations, *Stage* and *Connect*, highlight their “mission” as the starting point for determining their performance measurement “benchmarks”; these are subsequently translated into multiple goals, budgeting, and specific measurement dimensions.

“At the very top level, the social mission, which is both our organization mission and vision, distils down into four top-level goals and each goal a series of three objectives. And we measure how we have done against those at six monthly intervals” (Stage_Artistic Director)

Here the artistic director appears to be content with a rather rigid reporting framework which operates according to a prescribed format. For *Stage*, performance measurement is embedded within each operating segment, for example, programming, financing, marketing, and development. It is used primarily as a strategic management tool, to assess whether the goals and objectives set out in the organization’s business plans, reviewed every 2 years, are being met. In terms of reporting and disclosure, results are presented both to the board of directors, trustees, regulatory bodies, and funders.

While *Stage*’s approach to performance measurement is formalized and prescribed, due to its unpremeditated operating model, *Connect*’s approach is more relaxed and organic, aiming to identify suitable funding opportunistically. The director of *Connect*, however, expressed growing concerns due to a lack of the resources and expertise essential for performance measurement. She expressed her discomfort when reporting the impact of the organization’s activities in numeric terms and her preference for communicating by narratives.

“... certainly, we don’t have the resources ... we count numbers, we talk to people and hear what their experiences were like so that we can improve on the next one. It’s not formulated, or recorded, apart from probably on social media about events that we develop. At the end of the day you should be able to fill out the report form for the funder and say ‘look we had so many visitors or we got so many speakers, or we worked with these many partners based on the data; ... but there’s not a framework that I work to because every project we do is completely different ... and I don’t have a personal set of key performance indicators. I’m asked to produce documents to justify things all the time rather than being able to explain them in my roundabout way ...” (Connect_Director)

In emphasizing the uniqueness of each of the projects with which her organization is involved, *Connect*’s director signals a requirement for something more akin to a “*phronetic*” approach to “experiential” performance measurement rather than the more formulaic model with which the organization is required to work.

Both organizations, however, expressed their frustration in being unable to “prove” what they were doing and were eager to identify their societal impact more clearly. They both had plans to commission someone to measure this impact “properly” in order to “improve” their performance.

“I really would like to know how far and wide our impact is ... How do you measure how much you inspire people? ... I don't know how you could measure what effect you have on people and what else happens, because the only way to measure things is really with numbers and stuff.” (Connect_Director)

Connect's director, in highlighting the deficiencies of quantitatively focused reporting and its inability to capture the “contextually embedded” emotional impact of artistic productions, appears to be implicitly acknowledging the deficiencies of “*techne*” or “*episteme*” as paradigms for reporting on the arts. This is reinforced by comments made by *Stage's* finance manager.

“In this day and age of budgets and cost cutting, and data driven evaluation, we can lose some of the qualitative value we contribute ... what else I would like to gather ideally is about looking at our lives here, measuring things like happiness, learning, feeling better about life, about feeling connected with people, about how new ideas can germinate through the arts and excite and inspire people. We can lose sight of some of that as it's very hard to capture and evidence.” (Stage_Finance Manager)

In order to overcome the limitation of counting participants and asking them to tick boxes on an evaluation form, the artistic director at *Stage* experimented by giving a named individual a remit for tracking changes within a project.

“We are doing a long-term, across a number of years, equalities-based piece of work, where we are embedding an artist into that project to look at helping us understand the changes that we're making to document and record reflections on that. So that we can track the change that's happening at different stages within the project ... through more qualitative embedded approaches, we are trying to understand the complexity and the detail of a particular piece of performance.” (Stage_Artistic Director)

Stage's artistic director is acknowledging its inability to capture the intangible, “interactively emergent,” aspects of artistic performance. The artistic director suggested that pre and post evaluation forms or any linear form of reporting based on numbers could not capture “what is in the air.” The artistic director's endeavors were, however, viewed as “romanticism” by *Stage's* board chair. While the chair saw no point in challenging “the system,” the directors of both organizations were eager to make measurement more relevant to practice, motivated by the long-term development of their organizations.

“... A lot of effort can go into trying to design and measure the output ... Sometimes there is too much effort in that, I can draw a parallel with the local authority cutting the grass and cleaning the streets. Nobody measures the impact of cutting grass, there is no big report on the social, economic, educational, and other benefits of grass cutting. Some people require hard outputs so for grass cutting you can measure what height the grass is, how many times it has been cut. We can do the same, how many shows we put on, how many people come in but as for the true impact the organization has and the changes it makes to people's lives of this area

.... How do you measure the impact of anger, love, curiosity, surprise, shock, depression all of those things that a truly fantastical quality experience can have on an individual? We can disappear up our own backsides if we spend all of our time and effort trying to measure those sorts of impacts. Water is to fish as culture is to humans.” (Stage_The Chair of Board)

Given *Stage's* limited resources, the chair believed that providing information on what funders specifically ask for was sufficient or, at least, all that was possible.

“We can performance measure the staff in terms of their job roles and all of that sort of stuff and Creative Scotland can say this is what we gave you money for and are you delivering on this or that, they have their own measures that they use to determine that” (Stage_The Chair of Board)

Nevertheless, even here, a key internal stakeholder acutely aware of the reliance of his organization upon external funders and their demands for “relevant” reporting, points out the limitations of such reporting. In summary, although *Stage's* reporting procedures are more formalized than *Connect's*, with significantly greater levels of capacity and resources, internal stakeholders within both organizations appeared aware of the deficiencies inherent in the formal systems of external funding agencies they were required to report to.

5.2 | Measurement from the outside

The interviewee from the local council described their relationship with *Stage* and *Connect* as collaborators and partners as they hosted joint events, and the council, by not only providing funding but also a meeting place, orchestrates all arts organizations in the city to ensure that programs complement one another thus avoiding audience-splitting on any particular night. He indicated that the council would assess projects based on how well agreed objectives have been met. He also opined that to move forward, the council should find ways to support rather than merely handover cash and that more detailed conversations between different parties are necessary. He also regretted that, sometimes, the “detached, measurement and logic-based” evaluation that the council undertook was not as efficient as one might hope.

“We'll be looking at the number of performances they're putting on, how they were reaching out to perhaps 'hard to reach' audiences. But there's also trying to ensure they are offering a broad range of cultural activities. That goes beyond the mainstream. You would hope that you would get more independent cinema than blockbusters. So, they get bums on seats but they also get good ratings and so on. And I would also look at them to be financially sound and well managed, because all of the foregoing stuff is useless if they've gone bust, same as any organization.” (Local Council Staff)

Another local council officer, who has followed *Connect's* activities closely, discussed the council's difficulty when assessing the performance of different organizations. Limited resources and capacities are not only issues for arts organizations but also for their funders. The council and the national funding agency are constantly evaluating themselves with most of

their funding coming from the Scottish Government and the UK National Lottery which themselves are subject to austerity and uncertainty. When the local council representative was asked as to how much effort has been put into assessing and interpreting the performance measurement reports submitted by their grantees his response was rather skeptical.

“Very little actually to be perfectly honest. Not enough. The council's really small, it's a small rural council. The council's got very little resources and they've had a lot of staff cuts, so there isn't really enough capacity to be measuring individual performance of businesses, so they tend to look at the sector or economy as a whole.”
(Local Council Officer)

The local council officer expressed uncertainty about how to interpret, understand and use narratives given their subjective nature. The external stakeholders interviewed, however, were unanimous in stating it was important to portray a balanced picture of the sector's activities rather than one solely focused on “universal” or “nomothetic” variables which could be measured easily. Thus, not only internal, but also external stakeholders, were aware of the deficiencies inherent in the reporting systems with which they were familiar. The director of the national funding agency suggested that the reciprocal relationships between funders and receivers were not explicitly transactional and that judgments should be based on consensus among those assessing progress on projects.

“So, it's not about just measuring. It's actually about ensuring there is an integrity to process self-assessment and improvement that those organizations are engaged with ... The important thing is transparency around discourse so that there is realism, pragmatism around what might be possible, about where interventions can be seen to be or can be perceived to be having an effect, ensuring that perception is articulated and shared and that it's honest. If we see an effect that we might have expected for 100 people but it's only happening for four, it's happening in a really positive way for those four.” (National Funding Agency Director)

This director of the national funding agency, who had been trained as an artist for many years and worked as an art worker himself, strongly believed that funders should make judgments based on “intersubjective” dialogue with individual art organizations, rather than on what could be easily assessed.

“Because in the cultural sphere, you're often generating the values by which your significance is actually assessed. That's a really, really dangerous space. We have to be very, very careful that we just don't position ourselves at the very center of our own assessment and evaluation procedures. We need to have some kind of way of challenging and assessing that. That often comes through constructing a dialogue around the work, constructing challenge, ensuring that the challenges that are evolving as society develops a broad sense of that and that we can do that in a way that's obviously got an integrity attached to it. It becomes not an objective measure; it becomes the measure of ‘intersubjective’ dialogue.” (National Funding Agency Director)

In summary, privileging formalized reporting procedures based on *techne* or *episteme*, external stakeholders require transparency in reporting for accountability, yet at the same time warn

of the danger of assessing only what is (easily) “measurable.” Recognizing challenges in assessing artistic outcomes and social impacts, the national funding agency director favors constructing dialogue “around” the work rather than making judgments based only on “objective” measures. Both interviewees, those representing local and national funding bodies, however, articulated concerns regarding institutionalized funding systems which impose performance measurement and reporting based on formalized rules and procedures. Both struggle with limited resources and with performance measurement systems which fail to capture the essence and *raison d’être* of arts organizations.

5.3 | Tensions between the inside and the outside

The previous two sub-sections have demonstrated that tensions between internal and external stakeholders are less explicit than our a priori presumptions led us to expect might be the case. Comments by both sets of stakeholders highlight the deficiencies of formalized reporting grounded in “*techné*” or “*episteme*” and internal stakeholders were aware of externals’ needs. For example, *Stage*’s performance measurement is complex and sophisticated and based on a combination of evaluative principles including self, peer group, and stakeholder assessments. The finance manager emphasized that each report prepared showed the success indicators required by the target stakeholders.

“The funders stipulate the information required, therefore in submitting a bid, one has to be very clear about what can and will be provided, and then that has to be reported. So, it is an agreement between the funder and *Stage*” (Stage_Finance Manager)

Funders play an important role for both *Stage* and *Connect*. *Stage* also uses their “Year End Briefing,” one of their most informative reports, to influence existing and potential funders.

“This document will be going to anybody who’s funded or donated to us, council, university, heads of departments and schools, and MPs and MSPs. It will also be on our website so it’s public ... And we try and get it to people we think will be interested in it, who could influence, for example, funding for the arts.” (Stage_Admin and Development Manager)

Both organizations’ managers emphasized the importance of handling financial data rigorously and of being transparent.

“... by the very nature of the kind of work we do ... staff quite often inherently find paperwork challenging, but everything has to be written down. You do have to make sense of how you spent it, but you should also be trusted to spend it. We have a public, ethical responsibility to make sure that money has been well spent.” (Connect_Project Manager)

However, while both organizations aim to deliver projects as rigorously as possible and report on their accountability relating to funded projects, the director of *Connect* criticized funders for often being too rigid. The director stressed that presenting different stories, “like

speaking different languages,” was critical in order to cater for individual stakeholder needs.

“They were expecting us to act as a department of the council—‘you have to deliver X.’ And this led to resistance, and the conversation went upside down ... we got back around the table and started again ... it became clear that their approach was too rigid, and that wasn’t going to work ... so we had quite a confrontational conversation ... It’s necessary to have somebody who has got the ability to connect and talk all those different languages, not to corral the artists.” (Connect_Director)

About 40% of *Stage’s* revenue comes from the national funding agency, and initiatives are based on the shared interests of both organizations in community, social inclusion, and environment-related projects. The operations manager, however, indicated her concern about organizational engineering based on over-measurement and reporting for “accountability,” and shed light on misalignment between the positive social impact targeted by the organization, and frustratingly time-consuming reporting requirements based on collating quantitative data which defied meaningful interpretation.

“The comprehensive report has a really detailed breakdown ... things about all of our diversity monitoring statistics, environmental statistics, and that goes right down to the level of things like how many bin bags of rubbish we throw away versus how many we recycle. So really detailed nitty-gritty, some of which is related to our mission, some of which is more related to theirs. It’s not that it’s a problem to do it, but it is a focus of our energy that is more driven by them than driven by us. We are actually really interested in our environmental impact, but we would explore it and prioritize it differently to the way that their monitoring encourages. Things like, ‘How many sheets of paper did you print?’ or ‘How many light bulbs did you use?’ I think we’re more interested in bigger scale and more interesting environmental changes.” (Stage_Operations Manager)

She also called for more “democratic approaches” based on trust when reviewing and agreeing project expectations. The directors of both organizations continually reiterated the desirability of telling a “story” as part of the reporting process.

“... in terms of interpreting our performance for external audience, will be more about turning that into a form of narrative, a form of prose ... that’s really about getting people engaged in the story of what we are doing and how that’s important as much as other indicators like volume, the number of people that hears what the critic says. It’s actually ‘let’s tell you what we have been up to and within that you can form your own opinion on whether we are delivering on our mission as an organization ... You are telling them a story of what we set out to do and how we did it ... It’s ‘how’ you tell the story that’s as important as anything else ...” (Stage_Artistic Director)

Summarizing, we find that the perspectives of internal and external stakeholders differ in terms of a number of performance measurement dimensions and often their values, objectives, and expectations conflict. Resultant tensions are therefore inevitable. Both internal

and external stakeholders, however, are also “trapped” within institutionalized systems which requires out-of-context data-driven performance measurement while also setting arbitrary benchmarks. Internal stakeholders struggle not only with rigid and inflexible performance measurement systems which fail to offer viable alternatives to the status quo, but also with lack of trust and understanding on the part of funders at times. While reservations were also expressed about the deficiencies of formalized reporting systems, and their inability to capture much of the essence of the arts, such reservations were acknowledged by both internal and external stakeholders and there is evidence of positive interaction between both sets of stakeholders.

6 | DISCUSSION

The performance measurement practices of the selected arts organizations are conceptualized following the two paradigms introduced by Evered and Louis (1981, p. 385) which contrast perspectives from “the inside” and “the outside” in endeavoring to understand the tensions between internal and external stakeholders. Table 3 contrasts the two perspectives on a number of dimensions, including ways of knowing, relationship to the organization, validation basis, source of indicators, context, aim, type of knowledge, nature of data, meaning, purpose, communication, mechanisms and benchmarks (Evered & Louis, 1981).

We find that internal stakeholders are more interested in immediate audience experience, and in capturing more valid, useful, and relevant knowledge (“*phronesis*”) (Flyvbjerg, 2001; Polkinghorne, 2004). They measure from the inside without necessarily recognizing the validity

TABLE 3 Measurement from the inside and the outside.

Measurement dimension	From the inside	From the outside
Ways of knowing	<i>Phronesis</i>	<i>Episteme & Techné</i>
Relationship	Involved	Detached
Validation basis	Experiential	Measurement & logic
Source of indicators	Interactively emergent	A priori
Context	Situational relevance	Universality & generalizability
Aim	To improve	To prove
Type of knowledge	Particular & idiosyncratic	Universal & nomothetic
Nature of data	Interpreted	Factual
Meaning	Contextually embedded	Context-free
Purpose	Coping	Accountability
	Sense-making	Screening
	Surviving	Organizational engineering
Communication	Dialogue	Intersubjective dialogue
Mechanisms	Narrative	Narrative
	Figures	Figures
Benchmarks	Mission	Measurable objectives

Source: Adapted from Evered and Louis (1981).

of a priori indicators explicitly while being experientially involved in their organizations. The purposes of performance measurement are fostered by their features of coping, sense-making, and surviving mechanisms. External stakeholders, on the other hand, are more detached and more interested in measuring based on a priori indicators generalizable to other situations. They tend to follow a centralized and formalized preselected set of indicators for accountability which may lead to a form of perceptual “screening,” so that they see only “what is being sought” (“*episteme*” and “*techné*”).

The main difference between the two perspectives is the level of appreciation of the performance measurement context (Chiaravalloti & Piber, 2011; Gilhespy, 1999; Labaronne, 2017). Internal stakeholders are directly involved in the research setting and are able to interpret their performance in light of the “context,” while the needs of external stakeholders often necessitate stripping away the idiosyncrasies of the particular organization studied, and consequently, collecting data which are considered to be “context-free.”

Internal stakeholders stress the importance of “dialogue” when communicating with funders about project outcomes. However, the external stakeholders whom we interviewed also believed that, without the stories around numbers which provide context, they could not make appropriate judgments. Similarly, Schuster (1996, p. 266) argues that funding agencies should be concerned not only with the nature of performance indicators but also with their use. This is similar to the argument of ter Bogt and Tillema (2016) on enabling “open access” to accounting information and “loosening control”; such initiatives might ultimately foster trust between arts organizations and funding agencies and build long-term relationships. Sundström (2011) comments that reporting on subjective performance externally is complex, emotional and the perceived usefulness of such measurements is dependent on the distance and trust between information users as well as knowledge and understanding of relevant contextual factors.

The findings highlight the importance of “intersubjective” dialogue which emphasizes the shared understanding of cultural insights and challenges and does not aim to “objectify” measures and metrics but rather to ensure that arts organizations are conscious of the values and outcomes that they generate. According to Gillespie and Cornish (2010), “intersubjectivity” refers to agreement based on the common-sense, shared meanings constructed by people in their interactions with each other. As a response to inter-group confrontation and conflict, “intersubjectivity” emphasizes mutual orientation between representatives of institutions with different histories, constraints, interests, and perspectives. When both parties share a “subjective definition of the situation” or a “mutual awareness of agreement and the realization of such understanding,” this may lead to co-construction of values based on a reciprocal relationship.

6.1 | Theoretical contributions

Focusing specifically on arts organizations, our study provides nuanced arguments for why and how internal and external stakeholders perceive performance measurement practices differently and makes three contributions to the nonprofit performance measurement literature. First, previous studies on nonprofits stress the need to explore the tensions between different stakeholders (Carman, 2007, 2009; Carman & Fredericks, 2010), although few previous studies are concerned specifically with arts organizations. As is the case for other nonprofits, much of the tensions between internal and external stakeholders in arts organizations arise as a result of imposed institutionalized systems which privilege “universal” or “nomothetic” measures and which lack relevance to practice. Additionally, however, the predominantly expressive nature of the arts and the

belief systems of the internal stakeholders of arts organizations, mean that performance measurement in such organizations is more difficult than in other nonprofits. Given internal resistance to the ethos of “measurement,” arts organizations struggle to translate their artistic objectives and achievements into valid and reliable performance indicators. Many arts organizations are either micro or small-sized organizations while also being significantly dependent on government funding; this means that arts organizations are faced with a distinctive set of challenges and tensions between different stakeholders as compared with other nonprofits.

Second, our study provides empirical evidence as to how arts organizations in different contexts perceive performance measurement related challenges differently by comparing two different-sized arts organizations, especially in terms of levels of capacity and performance measurement resources. Prior positivist literature fails to address the contextual complexity of arts organizations, whereas our study involves in-depth exploration of the unique methods, techniques, and practices used as well as the ethics of each arts organization.

Third, in applying Evered and Louis's (1981) alternative paradigms of inquiry from “the inside” and “the outside,” our study achieves articulation of the epistemological differences between internal and external stakeholders' perspectives through the lens of practice theory (Flyvbjerg, 2001; Polkinghorne, 2004). Such an appreciation is a prerequisite to developing appropriate performance measurement approaches encapsulating tolerance, diversity and multiplicity and focusing less on technical “correctness.”

6.2 | Managerial implications

Our study offers implications for managerial practice. First, we emphasize the lack of capacity and resources experienced by arts organizations and concerns around the effort and time expended by internal stakeholders in collating performance measurement data in an environment of constant funding cuts (Benjamin, 2010; Carman & Fredericks, 2010; Weinstein & Bukovinsky, 2009). Overtly complicated performance measurement criteria create organizational problems and tensions; thus, it is critical to simplify the criteria used. Arts organizations should use their limited resources more strategically (Carman, 2007, 2009; Carman et al., 2008). Funders should also consider providing accessible training to enhance evaluation capacity or aim to develop “hands-on” relationships through greater investment of time, money, and expertise, so that arts organizations are able to formulate achievable strategies, benchmarks, and performance measures (Benjamin, 2010; Cobb, 2002; Ostrander, 2007). Second, our study highlights the limitations of formalized reporting systems to capture the specific artistic and social impacts of arts organizations. Echoing Labaronne and Tröndle (2021), both internal and external stakeholders in our study appear to acknowledge a need to ground performance assessment in artistic and social value and there are indications of a degree of positive engagement between both sets of stakeholders (Labaronne, 2017). Our study identifies increasing efforts by local and national funders to take account of narratives, as well as numbers, based on more “holistic” systems. These stakeholders emphasize communication based on a “dialogue” as well as a shift to a more long-term focus. We recommend funders to be tolerant to more “imaginative” approaches rather than conforming purely to prescribed, or more “literal” measures and approaches to accountability (Campbell, 2002). Finally, we recommend arts organizations to become more adept at “measuring” artistic quality and its consequential impact longitudinally in accordance with external demands which determine the criteria upon which they are adjudicated.

7 | CONCLUSIONS

By investigating epistemological differences between internal and external stakeholders' perspectives on performance measurement by arts organizations, our article fills a lacuna in current discourse. Although the two arts organizations we studied differ with regards to managerial skill sets, resources, and capacities, both approach performance measurement seriously and rigorously. *Stage*, with established systems based on the use of various indicators for each measurement dimension identified, is more confident than *Connect* that measurement provides the organization with useful information about its overall performance. The study highlights a persisting paradox in performance measurement and reporting since the internal stakeholders of arts organizations are under great pressure to measure everything "which can be measured," while the essence of arts organizations' performance, as articulated in their mission statements, remains immeasurable.

Arts organizations should not alter their values to align with funders or external constraints but should "embrace" their mission and exploit external opportunities aligned with their values. At the same time, funding agencies should progress from "controlling" measurement toward trust in the efficacy of collaboration based on shared interests, privileging "responsiveness" rather than "efficiency." Based on this notion, managers in the arts and cultural sector may be usefully encouraged to invest more time in "improving" rather than "proving" their performance. There are grounds for optimism in that external stakeholders, as well as those "on the inside," are aware of the deficiencies of performance measurement systems which fail to engage with the artistic activity which is the *raison d'être* of arts organizations.

Our research makes novel contributions in terms of advancing existing discourse around the tensions between nonprofit stakeholders by utilizing Evered and Louis's (1981) inquiry from "the inside" and "the outside" paradigms and articulating the epistemological differences between internal and external stakeholders' perspectives through the lens of practice theory (Flyvbjerg, 2001; Polkinghorne, 2004). Nevertheless, it does have limitations which create avenues for future research. Our case study arts organizations are located within an Anglo-Saxon country and are funded primarily by government subsidies; thus, they may be more predisposed to interact positively with external demands for evaluation. Arts organizations in non-Anglo-Saxon countries may be prejudiced against performance measurement practices (Labaronne, 2017). Consequently, our findings may not be generalizable to non-Anglo-Saxon countries or countries with different funding systems for arts organizations and should be interpreted with caution.

ACKNOWLEDGEMENT

Open access publishing facilitated by University of South Australia, as part of the Wiley - University of South Australia agreement via the Council of Australian University Librarians.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Boram Lee  <https://orcid.org/0000-0002-1291-7188>

REFERENCES

- Abdullah, A., Khadaroo, I., & Napier, C. J. (2018). Managing the performance of arts organisations: Pursuing heterogeneous objectives in an era of austerity. *The British Accounting Review*, *50*(2), 174–184.
- Atkinson, A. A., Waterhouse, J. H., & Wells, R. B. (1997). A stakeholder approach to strategic performance measurement. *MIT Sloan Management Review*, *38*(3), 25–37.
- Balser, D., & McClusky, J. (2005). Managing stakeholder relationships and nonprofit organization effectiveness. *Nonprofit Management and Leadership*, *15*(3), 295–315.
- Becker, K., Antuar, N., & Everett, C. (2011). Implementing an employee performance management system in a nonprofit organization. *Nonprofit Management and Leadership*, *21*(3), 255–271.
- Benjamin, L. M. (2010). Funders as principals: Performance measurement in philanthropic relationships. *Nonprofit Management and Leadership*, *20*(4), 383–403.
- Boerner, S. (2004). Artistic quality in an opera company: Toward the development of a concept. *Nonprofit Management and Leadership*, *14*(4), 425–436.
- Boorsma, M., & Chiaravalloti, F. (2010). Arts marketing performance: An artistic-mission-led approach to evaluation. *The Journal of Arts Management, Law, and Society*, *40*(4), 297–317.
- Campbell, D. (2002). Outcomes assessment and the paradox of nonprofit accountability. *Nonprofit Management and Leadership*, *12*(3), 243–259.
- Carman, J. G. (2007). Evaluation practice among community-based organizations: Research into the reality. *American Journal of Evaluation*, *28*(1), 60–75.
- Carman, J. G. (2009). Nonprofits, funders, and evaluation: Accountability in action. *The American Review of Public Administration*, *39*(4), 374–390.
- Carman, J. G., & Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation*, *31*(1), 84–104.
- Carman, J. G., Fredericks, K. A., & Introcaso, D. (2008). Government and accountability: Paving the way for nonprofits and evaluation. *New Directions for Evaluation*, *2008*(119), 5–12.
- Cavalluzzo, K. S., & Ittner, C. D. (2004). Implementing performance measurement innovations: Evidence from government. *Accounting, Organizations and Society*, *29*(3–4), 243–267.
- Chiaravalloti, F. (2014). Performance evaluation in the arts and cultural sector: A story of accounting at its margins. *The Journal of Arts Management, Law, and Society*, *44*(2), 61–89.
- Chiaravalloti, F., & Piber, M. (2011). Ethical implications of methodological settings in arts management research: The case of performance evaluation. *The Journal of Arts Management, Law, and Society*, *41*(4), 240–266.
- Cobb, N. K. (2002). The new philanthropy: Its impact on funding arts and culture. *The Journal of Arts Management, Law, and Society*, *32*(2), 125–143.
- Cochran, D., David, F., & Gibson, C. K. (2008). A framework for developing an effective mission statement. *Journal of Business Strategies*, *25*(2), 27–39.
- Donovan, C., & O'Brien, D. (2016). Governing culture: Legislators, interpreters and accountants. *Critical Perspectives on Accounting*, *37*, 24–34.
- Eikenberry, A. M., & Kluver, J. D. (2004). The marketization of the nonprofit sector: Civil society at risk? *Public Administration Review*, *64*(2), 132–140.
- Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25–32.
- Evans, G. (2000). Measure for measure: Evaluating performance and the arts organisation. *Studies in Cultures, Organizations and Societies*, *6*(2), 243–266.
- Evered, R., & Louis, M. R. (1981). Alternative perspectives in the organizational sciences: “Inquiry from the inside” and “inquiry from the outside”. *Academy of Management Review*, *6*(3), 385–395.
- Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge University Press.
- Froelich, K. A. (1999). Diversification of revenue strategies: Evolving resource dependence in nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, *28*(3), 246–268.
- Frumkin, P. (2002). *On being nonprofit: A conceptual and policy primer*. Harvard University Press.
- Garengo, P., Biazzo, S., & Bititci, U. S. (2005). Performance measurement systems in SMEs: A review for a research agenda. *International Journal of Management Reviews*, *7*(1), 25–47.

- Gilhespy, I. (1999). Measuring the performance of cultural organizations: A model. *International Journal of Arts Management*, 2(1), 38–52.
- Gillespie, A., & Cornish, F. (2010). Intersubjectivity: Towards a dialogical analysis. *Journal for the Theory of Social Behaviour*, 40(1), 19–46.
- Gilmore, A., Glow, H., & Johanson, K. (2017). Accounting for quality: Arts evaluation, public value and the case of “culture counts”. *Cultural Trends*, 26(4), 282–294.
- Herman, R. D., & Renz, D. O. (1997). Multiple constituencies and the social construction of nonprofit organization effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 26(2), 185–206.
- Herman, R. D., & Renz, D. O. (1998). Nonprofit organizational effectiveness: Contrasts between especially effective and less effective organizations. *Nonprofit Management and Leadership*, 9(1), 23–38.
- Herman, R. D., & Renz, D. O. (1999). Theses on nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 28(2), 107–126.
- Herman, R. D., & Renz, D. O. (2004). Doing things right: Effectiveness in local nonprofit organizations, a panel study. *Public Administration Review*, 64(6), 694–704.
- Johnston, R., & Pongatichat, P. (2008). Managing the tension between performance measurement and strategy: Coping strategies. *International Journal of Operations & Production Management*, 28(10), 941–967.
- Kaplan, R. S. (2001). Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management and Leadership*, 11(3), 353–370.
- Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, 74(1), 75–85.
- Krug, K., & Weinberg, C. B. (2004). Mission, money, and merit: Strategic decision making by nonprofit managers. *Nonprofit Management and Leadership*, 14(3), 325–342.
- Labaronne, L. (2017). Performance measurement and evaluation in arts management. *Zeitschrift für Kulturmanagement*, 3(1), 37–70.
- Labaronne, L., & Piber, M. (2020). Performance measurement and evaluation in the arts and cultural sector: State-of-the-art in theory and practice and prolegomena for further developments. In *Management, participation and entrepreneurship in the cultural and creative sector* (pp. 219–240). Springer.
- Labaronne, L., & Tröndle, M. (2021). Managing and evaluating the performing arts: Value creation through resource transformation. *The Journal of Arts Management, Law, and Society*, 51(1), 3–18.
- Lecy, J. D., Schmitz, H. P., & Swedlund, H. (2012). Non-governmental and not-for-profit organizational effectiveness: A modern synthesis. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 23(2), 434–457.
- Levy, P. E., & Williams, J. R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*, 30(6), 881–905.
- Meyrick, J., Phiddian, R., & Barnett, T. (2018). *What matters? Talking value in Australian culture*. Monash University.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook*. Sage Publications.
- Moxham, C. (2010). Help or hindrance? Examining the role of performance measurement in UK nonprofit organizations. *Public Performance & Management Review*, 33(3), 342–354.
- Nørreklit, H. (2011). The art of managing individuality. *Qualitative Research in Accounting & Management*, 8(3), 265–291.
- Ostrander, S. A. (2007). The growth of donor control: Revisiting the social relations of philanthropy. *Nonprofit and Voluntary Sector Quarterly*, 36(2), 356–372.
- Pandey, S., Kim, M., & Pandey, S. K. (2017). Do mission statements matter for nonprofit performance? Insights from a study of US performing arts organizations. *Nonprofit Management and Leadership*, 27(3), 389–410.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage Publications.
- Peterson, P. A. (1986). From impresario to arts administrator. In P. DiMaggio (Ed.), *Nonprofit enterprise and the arts* (pp. 161–183). Oxford University Press.
- Phiddian, R., Meyrick, J., Barnett, T., & Maltby, R. (2017). Counting culture to death: An Australian perspective on culture counts and quality metrics. *Cultural Trends*, 26(2), 174–180.
- Polkinghorne, D. (2004). *Practice and the human sciences: The case for a judgment based practice of care*. Suny Press.

- Rentschler, R., Fillis, I., & Lee, B. (2022). Artists versus arts council: A longitudinal analysis of brand legitimacy. *Poetics*, 92, 101623.
- Rentschler, R., Lee, B., & Fillis, I. (2021). Towards an integrative framework for arts governance. *International Journal of Arts Management*, 24(1), 17–31.
- Rentschler, R., Lee, B., & Subramaniam, N. (2021). Calculative practices and socio-political tensions: A historical analysis of entertainment, arts and accounting in a government agency. *Accounting History*, 26(1), 80–101.
- Sawhill, J. C., & Williamson, D. (2001). Mission impossible?: Measuring success in nonprofit organizations. *Nonprofit Management and Leadership*, 11(3), 371–386.
- Schuster, J. M. (1996). The performance of performance indicators in the arts. *Nonprofit Management and Leadership*, 7(3), 253–269.
- Sundström, A. (2011). Framing numbers “at a distance”: Intangible performance reporting in a theater. *Journal of Human Resource Costing & Accounting*, 15(4), 260–278.
- ter Bogt, H., & Tillema, S. (2016). Accounting for trust and control: Public sector partnerships in the arts. *Critical Perspectives on Accounting*, 37, 5–23.
- Throsby, D. (2017). Defence of measurement: A comment on Phiddian et al. *Cultural Trends*, 26(4), 314–317.
- Turbide, J., & Laurin, C. (2009). Performance measurement in the arts sector: The case of the performing arts. *International Journal of Arts Management*, 11(2), 56–70.
- Voss, G. B., Cable, D. M., & Voss, Z. G. (2000). Linking organizational values to relationships with external constituents: A study of nonprofit professional theatres. *Organization Science*, 11(3), 330–347.
- Voss, Z. G., & Voss, G. B. (2000). Exploring the impact of organizational values and strategic orientation on performance in not-for-profit professional theatre. *International Journal of Arts Management*, 3(1), 62–76.
- Weinstein, L., & Bukovinsky, D. (2009). Use of the balanced scorecard and performance metrics to achieve operational and strategic alignment in arts and culture not-for-profits. *International Journal of Arts Management*, 11(2), 42–55.
- Yin, R. K. (2015). *Qualitative research from start to finish*. Guilford Publications.

AUTHOR BIOGRAPHIES

Boram Lee is a Senior Lecturer in Arts and Cultural Management, University of South Australia. She is specialised in the field of accounting and finance with an emphasis on behavioural studies based on psychological approaches. Boram also has a wide range of research interests in cross-cultural and cross-disciplinary studies, covering the valuation of arts and culture, creative fundraising, governance and accountability, artists' labour markets and entrepreneurship.

Rodolfo Sejas-Portillo is currently researching the redistributive effects of information disclosure policies. His previous research has studied the effects of psychological biases in functioning markets.

Ian Fraser is an Emeritus Professor of Accounting at the University of Stirling. He has also held academic posts at the University of Strathclyde and Glasgow Caledonian University. Ian's primary research interests are in auditing, financial reporting and corporate governance and he has published extensively in these areas. He also has an interest in the valuation of art and artworks and similar forms of capital.

How to cite this article: Lee, B., Sejas-Portillo, R., & Fraser, I. (2023). Balancing perspectives on performance: “Measurement from the inside” and “measurement from the outside”. *Nonprofit Management and Leadership*, 1–22. <https://doi.org/10.1002/nml.21561>