

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/16596>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

The Paradox of Impact Measurement in Cultural Contexts

Andrew Anzel,¹ Haley Beer² & Graeme Currie³

ABSTRACT

The cultural sector has increasingly adopted practices of impact measurement to justify public investment. While scholars have investigated the (de)merits of this cultural trend, literature on the rationale, development, and evolution of impact measurement within cultural contexts is limited. We address this oversight by uncovering the intellectual history of impact measurement in cultural contexts. We reveal how measurement justification and practice became misaligned over time, creating a latent paradox of impact measurement in cultural contexts. We argue that current practices of impact measurement both strengthen and weaken the justification for public cultural investment. In the short term, impact measurement strengthens investment justification by providing cursory evidence of produced outcomes. In the long term, current practices weaken this justification by failing to produce counterfactual accounts for how cultural spend achieves funding objectives. We problematize this paradox and outline future research to either *confront*, *embrace*, or *cope with* its existence.

Keywords

Impact measurement, cultural investment, paradox, new public management

¹ Warwick Business School, UK. Corresponding author: A.Anzel@Warwick.ac.uk

² Associate Professor, Warwick Business School, UK.

³ Professor, Warwick Business School, UK.

The present policy, therefore, is not only 'unjust', but also sows the seeds of its own demise, for it is difficult to imagine that a major expansion will be sanctioned in expenditure in the Arts if the present policy continues – A.T. Peacock (1969).

INTRODUCTION

The application of impact measurement in cultural contexts has greatly expanded over the past few decades. Increasingly tied to public and private investment, impact measurement is said to enable the cultural sector to understand and act towards its theorized value creation (Throsby, 2017). For cultural practitioners, impact measurement is seen as a way for artists and cultural organizations to represent the noteworthy outcomes produced by their continued operation (Maliekel et. al., 2019). For cultural theorists, impact measurement renders visible the ways in which the cultural sector leads to economic regeneration (Seaman, 2020), social equity (Soh, 2019), and other instrumental impacts that justify the use of investment (Hadley & Gray, 2017).

While significant attention has been paid to refine the technologies of measurement (e.g., Maliekel et. al., 2019), scholarship on the rationale, development, and evolution of impact measurement within the cultural sector is limited. Existing literature currently connects the rise of impact measurement to broad policy perspectives, such as New Public Management (NPM) (Belfiore, 2004), and wider social trends, such as neoliberalism (Alexander, 2018). While both streams provide key insights into the rationalizing pressures that gave rise to impact measurement, scholarship has yet to provide a comprehensive examination of how additional pressures effected the evolution of measurement practices over time. Since impact measurement continues to be required by prominent cultural funding bodies (e.g., Arts Council England, 2021; Arts & Humanities Research Council, 2019), it is imperative to document and understand its intellectual history.

To address this gap, we reviewed extant scholarship to account for the rationale, development, and evolution of impact measurement within the UK cultural sector. We focused on the UK context due to the breadth of impact measurement scholarship previously published there, which allowed us to adopt a longitudinal lens for how the trend developed over time. Our review uncovers inconsistencies between the aims of using impact measurement to provide a rationalized justification for cultural investment, and what occurs in practice. For example, we show how current practices ignore tenets of NPM, such as counterfactual validation which

shows that the desired outputs could have been produced with less investment. We call this a 'justification-practice mismatch,' and argue it has created a latent paradox of impact measurement in cultural contexts.

While this paradox is developed over the course of our article, a brief sketch is warranted here. Impact measurement is theorized as a way to strengthen the justification for public cultural investment; however, current measurement practices *both strengthen and weaken* this justification. Current practices strengthen the case by providing cursory evidence of produced outcomes which may satisfy current funding requirements (e.g., n number of participants attended a weekly arts-for-wellbeing drawing course). Current practices weaken the case by failing to produce counterfactual accounts for how cultural spend achieves funding objectives (e.g., a weekly luncheon reached n² more people and made people happier, all the while costing the funder the same as the drawing course). We characterize this paradox as latent – the undermining effects of current practices are not felt in the present; however, future shifts in the funding environment might reveal the ill-supported evidence used to prop-up cultural investment, thus risking a funding collapse. To echo the words of our epigraph (a quote referring to early arts subsidy justification) the cultural sector may in fact 'sow the seeds of its own demise.'

In addition to providing an intellectual history of impact measurement in cultural contexts, our discussion of the found paradox offers further contributions to cultural scholarship. First, we note how current practices of impact measurement often use NPM discourse to mask the extension of historical investment justifications. Second, we discuss how this practice-justification misalignment may lead to the eventually delegitimization of public funding for the cultural sector if funders refocus on robust, demonstratable impacts. Third, we speculate that this misalignment may be an unintentional byproduct of the difficulty in representing the means-ends relationship between cultural interventions and their value creation. This difficulty creates variance between how accurately cultural contexts use impact measurement techniques, thus casting doubt on the utilization of impact measurement as a decision-making instrument to rationalize investment against achieving instrumental ends. Finally, we put forward a research agenda for cultural scholars and practitioners to either *confront*, *embrace*, or *cope with* the paradox of impact measurement in cultural contexts.

ANTECEDENTS OF IMPACT MEASUREMENT

Impact measurement is a relatively new phenomenon in cultural contexts, but this practice did not arise overnight. To understand the antecedents of impact measurement, we reviewed how public cultural funding has been governed and justified over time. To begin our review, we queried multiple academic databases for articles with key words related to ‘impact measurement,’ ‘culture,’ and ‘the arts.’ A large swatch of generated literature drew empirically from the UK context or referred to UK cases as indicative of wider trends. We dovetail our review within this trend and use the UK context to trace the antecedents, development, and evolution of impact measurement in cultural contexts. To support our analysis, we also reviewed grey literature, such as practitioner reports and funder strategy documents, that were referenced within our found academic literature. Our review uncovered that early justification for cultural investment was based on the positive externalities that arose from the creation and consumption of culture; however, mounting political critiques and pressure to ‘rationalize public spend’ weakened this early justification and led to the growth of impact measurement.

Early Justification for Cultural Investment

Explicit justification for public cultural investment is historically sparse. The first known justification for the purchase of art to be exhibited in the British Museum is put forward in a parliamentary proceeding, where the purchase of the Elgin Marbles was discussed in 1816, stating ‘the positive effect that the beautiful sculptures would undoubtedly have on the refinement of national taste’ (quoted in Belfiore, 2004, p.186). When the British Broadcasting Corporation (BBC) was founded in 1922, it quickly adopted the mission to ‘inform, educate and entertain’ (Scannell & Cardiff, 1991). John Reith, the founding General Manager, elaborating on this mission, stating that ‘our responsibility is to carry into the greatest possible number of homes everything that is best in every department of human knowledge... and to avoid the things which are, or may be, hurtful’ (Reith, 1924, p.34).

Another example of explicit justification can be found within the establishment of the Committee for the Encouragement of Music and the Arts (CEMA). CEMA was founded at the start of WWII in 1939 by the Pilgrim Trust in order ‘to prevent cultural deprivation on the home front’ (Netzer, 1978, p.197). The UK government assumed the reigns on CEMA in 1942 under the leadership of John Maynard Keynes, a famous economist and member of the Bloomsbury Group (Upchurch, 2004). CEMA was then reinvented into the Arts Council of

Great Britain. The Arts Council received its first royal charter in 1946, which described its objectives as:

To develop and improve the knowledge, understanding, and practice of the fine arts; to increase the accessibility of the arts to the public throughout Great Britain; and to advise and cooperate with Government Departments, local authorities, and other bodies on any matters concerned whether directly or indirectly with the foregoing objects (Quoted in Netzer, 1978, p.198).

While nomenclature often varied, the more disparate funding justifications given for cultural public spend centered around the positive externalities of culture. Externalities refer to the ‘spillover’ effects of the production and consumption of goods or services (Throsby, 2010). Midcentury economists posited that the arts were not only beneficial to those who ‘participated,’ but engendered a ‘more civilized’ society that everyone enjoyed (Reid, 2009). Evidencing this point, the economist Alan Peacock wrote that ‘the benefit is not merely discriminate, the positive effects of the fostering of art and learning and the preservation of culture are not restricted to those immediately prepared to pay cash but diffuse themselves to the benefit of much wider sections of the community in much the same way as the benefits of the apparatus of public hygiene or of a well-planned urban landscape’ (1969/1976, p.156). To use a contemporary example, scholars have reasoned that cultural investment may lead to national prestige, resulting in the positive externality of tourism income (Seaman, 2020).

Justification for cultural public spend based on positive externalities valued culture for its ‘civilizing, uplifting, and ennobling’ qualities (Reid, 2009). Evidencing this point, Keynes stated on a radio broadcast in 1945 that ‘the artist walks where the breath of the spirit blows him. He cannot be told his direction; he does not know it himself but leads the rest of us into fresh pastures and teaches us to love and enjoy what we often begin by rejecting, thus enlarging our sensibility and purifying our instincts’ (quoted in Netzer, 1978, p.203). Evidencing the perennial nature of this ennobling discourse is a note from Sir William Rees-Mogg, a later chairman of the Arts Council, that reads ‘on the whole I prefer to argue the case for funding the Arts Council on the contribution of the arts to British civilization’ (Rees-Mogg, 1984, p.V). As if the direct language was not enough, this quote appears in the forward to the 1984 strategic plan of the Arts Council entitled ‘The Glory of the Garden,’ a further reference to the cultivating nature of the arts.

Moving past its justification, midcentury public investment mainly originated from the Arts Council. The Arts Council was structured to distribute its monies according to the ‘arms-length principle.’ The arms-length principle characterizes the way in which quangos (quasi-autonomous non-governmental organization) operate (Quinn, 1997). Such organizations receive central funding from public bodies (e.g., central or regional governments) to achieve a given aim, though day-to-day operations are left solely to the organization (Hood, 2000). Founded as a quango, the Arts Council was funded through the Department of Education and Science but was originally funded with ‘no strings attached’ (Upchurch, 2004; Upchurch, 2016). The Arts Council would pass along this stringless investment to artists and arts organizations (Hutchinson, 1982). This allowed the Arts Council and its beneficiaries to make decisions based on taste (Throsby and Withers, 1979).

The way in which cultural investment was justified and applied within the UK was not without its critics. Turning first to its justification, Peacock (1969/1976) outlined the critique that ‘one cannot make a case on grounds of externality in absolute terms alone, however much one shares the value judgements of its proponents’ (p.77). This is because justifying the arts based on externalities creates ambiguity about ‘who’ should get funded or ‘to what extent’ they should be funded (Reid, 2009). Giving a more pointed critique, Peacock (1969/1976) questions whether current practices are income redistribution in reverse – subsidizing the habits of the rich while expecting benefits to the poor down the line. To avoid this problem, Peacock advocated for a more rational approach to cultural investment, stating that ‘one has to prove that cultural inputs, however provided, are more “productive” than any alternative way of attracting visitors using the same amount of resources. It is only when we assume that given inputs to culture the most efficient method that we arrive at a possible reason are for regional government subsidization’ (1969/1976, p.76).

Similarly, critiques emerged about how the Arts Council applied this justification to its funding decisions. Exemplifying this critique are King and Blaug (1976/1979), who examined decades of Arts Council reports to question its objectives and accounting practices. For example, they found that the Arts Council wrote that ‘there is a need to satisfy our readers who pay taxes but may not like reading balance-sheets, that we have not wasted their money and intend to spend it to even better advantage in the future’ (quoted in King & Blaug, 1976/1979, p.102). However, King and Blaug question ‘what is a “waste” of money in the field of the Arts,

and how would one recognise waste if one saw it?' (1976/1979, p.102). Taking this point further, they argue that:

'Even if we succeeded in drawing up a list of aims - objectives, goals, ends, the reasons why - we still face the task of somehow measuring the effectiveness of various spending patterns in achieving these aims. ...We label this evaluation technique 'cost-effectiveness analysis' and in principle it is as applicable to the arts as it is to any other area of government activity, although in practice it may be particularly difficult to apply to the Arts' (King & Blaug, 1976/1979, p.101).

This sentiment seems to have been recognized within the Arts Council itself, with one report stating, 'we shall try to make our judgements more scientific and less rule-of-thumb, but to a large extent they must remain inspirations - others might find a less flattering word' (quoted in King & Blaug, 1976/1979, p.118).

Amid these critiques and aspirations to become 'more scientific,' a wider change in how central governments operated also effected the continued development of cultural investment.

'Remaking' Investment Justification

While critics began to question the midcentury justification for public cultural investment, rationalization pressures on how the state supported all funding relationships fundamentally changed in the political upheaval of the 1980's with the advent of Thatcherism. In its aim to 'roll back the frontiers of the state,' the Conservative Government sought to reduce the open checkbook of the public purse and rationalize where it directed its spending. For example, the Financial Management Initiative was passed in 1982, which 'required each government department – as well as their subordinate agencies – to clearly identify their objectives and set targets against which their performance could be measured' (Belfiore, 2004, p.191). These reporting measures would then be used to judge whether claims on public resources could be considered legitimate (Jenkins 1995).

Thatcherism ebbed as time progressed, but its emphasis on rational public management served as a 'policy blueprint' for the rise of New Public Management (NPM). Theorized by Hood (1991), NPM characterizes the way in which public administration has evolved in light of principles drawn from economic theories of public choice, transaction cost, and principal-

agent, as well as other models of ‘scientific management’ (p.5). Moreover, its key beliefs are: cost control, financial transparency, the introduction of market mechanisms into the provision of public services, the reliance on a “contract culture” (Belfiore, 2004, p.191) and ‘the enhancement of accountability to customers for the quality of service via the creation of performance indicators’ (Power 1997, p. 43; see also Hood, 1991). This approach to public administration was further justified by management scholars who lauded impact measurement to mitigate the bounded rationality of leaders and decision-makers (Simon, 1997). While a full account of NPM’s history is outside the scope of this article, it has been well covered by Christensen and Lægheid (2016) and applied to the cultural context by Belfiore (2004).

What *is* important, is how NPM and its emphasis of impact measurement influenced public cultural investment. Demonstrating this point clearly is the government’s ‘Comprehensive Spending Review’ of the 1990’s. This review required the Department for Culture, Media, and Sport [DCMS] (the new departmental overseer of the Arts Council and other cultural spend) to set clear outcomes that were expected from its investment portfolio. As a result, DCMS would write in its 1998 Annual Report that ‘this is not something for nothing. We want to see measurable outcomes for the investment which is being made. From now on there will be real partnership with obligations and responsibilities.’ Furthermore, ‘DCMS ties its expenditure to its objectives’ and it needs ‘to be assured that public money is being used appropriately to meet public objectives,’ and with ‘more money invested, more results are required’ (DCMS quoted in Selwood, 2002, p. 12).

In tracing the genealogy between public support for culture and its policy justification within the UK context, we put forward that there were two main antecedents to the development of impact measurement in cultural contexts. First, early justification for cultural investment relied on arguments related to positive externalities – the arts create spillover effects for the whole population, not just those who create or consume them. However, critics questioned this underlying justification (e.g., income distribution in reverse) and argued that it was a poor guide for investment decisions. Second, political trends to become ‘more scientific’ resulted in policy that required recipients of public investment to account for their impact.. While we have thus far discussed the antecedents of impact measurement, we will now discuss how the cultural sector responded to these new institutional arrangements.

PRACTICES OF IMPACT MEASUREMENT

The coalescence of public funding for culture under the aegis of DCMS required the cultural sector to find new ways to account for and justify their receipt of public financial support. This process first required individual cultural organizations to explicitly state what theoretical impact they sought to create through their operation. In other words, they were required to make the tacit understanding underlying their work into textual, reportable forms. For example, until they were required to do so under the newly reconfigured DCMS, the British Museum published neither an annual report nor a corporate plan until 1996 (Edwards, 1996).

As more cultural organizations began to articulate their *raison d'être*, they began to report on easy-to-assess key performance indicators (Peacock, 2000). Such reporting consisted of measures related to ‘audience reach’ – demographics of audience, audience satisfaction, educational partnerships, as well as general revenue and expenses (DCMS, 2008). As time continued, DCMS also required specific figures related to funded organizations, such as staff demographics and an itemization of financial expenditure (Arts Council England, 2016b). While these general measures are still required by DCMS, cultural organizations developed more specific measures to further cement their argument of public support through appeals of impact.

Economic Impact Measurement

Cultural organizations first turned to the economist’s toolbox. Since DCMS had adopted an economic discourse, so too did cultural organizations seek to clarify what specific economic return they were producing with this investment. As such, an explosion of economic impact assessments of the cultural sector occurred in the 1980s and 1990s (Throsby, 2010). The publication widely regarded as the first to explicitly link economic impact with public investment was *The Economic Importance of the Arts in Britain* (Myerscough, 1988). This watershed document contained the findings of a gross value-added (GVA) study which attempted to capture the overall financial impact of the cultural sector on local, regional, and national levels. Since its publication, economic impact studies using GVA techniques have become a staple for the cultural field, with a similar report being published by Arts Council England as recently as 2019 (Arts Council England, 2019).

Using economic mechanisms to account for the value of cultural organizations has garnered both advocates and cynics within the cultural sector. Advocates claim such studies

‘portray public arts dollars as worthy investments, not subsidies or entitlements’ (Barsdate, 1997, p.18) and as a ‘most potent’ tool, ‘the one that resonates most with elected leaders, public and private sector funders, and policymaker’ (Cohen, Schaffer & Davidson, 2003, p.17). Conversely, cynics argue that ‘financial objectives are not everywhere harmonious with artistic or cultural objectives, so government intervention will at best be only coincidentally consistent with cultural objectives’ (Madden, 2001, p.172).

Social Impact Measurement

The cultural sector sought to answer calls to better understand the social dimensions of its impact and account for this additional value creation (Belfiore & Bennett, 2007; Galloway, 2009). Early studies began to link cultural participation with various improvements to a participant’s life. For example, it was noted that participation in cultural activities led to an improved self-concept and sense of control over one’s life (Lynch & Chosa, 1996). Others noted that individuals involved in creating or organizing cultural activity may demonstrate greater creativity and develop new artistic skills (Guetzkow, 2002). And yet others focused on culture’s role in education attainment (Fiske 1999), improving health (Costello, 1998), increasing social cohesion (Matarasso, 1997), and overall wellbeing (Taylor et al., 2015; Peaker & Vincent, 1990).

An exemplary case of measuring the social impact of a cultural program is that of New Victory Theater’s SPARK initiative (Maliekel et al., 2019). To address socio-cultural inequalities, SPARK was created as a multi-year partnership between an arts organization and schools in low-income neighborhoods to use artistic engagement to increase the social and creative development of students. To do this, SPARK schools were given free trips to performances, in-depth artistic residencies, and creative professional development for teachers. To account for their impact, the New Victory Theater performed a four-year, case control study of its partnership schools in comparison with similar schools who only received free trips to its performances. When compared with their control counterparts, researchers found that students in the SPARK intervention were able to score significantly higher on measures aimed to capturing emotional intelligence, empathy, and creative thinking in addition to performing better on their traditional academic subjects (Maliekel et al., 2019).

Social impact measurement techniques that have gained significant prominence in the cultural sector are under the header of theory-based evaluations, such as logic models or theory

of change models (Galloway, 2009). Such models are premised on creating ‘an explicit theory or model of how the program causes the intended or observed impact’ (Coryn, Noakes, Westine, & Schroter, 2011, p.201). Theory-based models seek to define causal pathways which link organizational inputs (e.g., strategic resources, capacities, associations) to envisioned activities (e.g., arts events, community development, lobbying) that would result in the creation of direct outputs which would benefit a local population. These immediate outputs were then expected to result in longer term outcomes for the community over time.

A major cultural organization to exemplify the use of a theory-based evaluation model is the Hull UK City of Culture Ltd (Hull Ltd.). Outlining their model within their final evaluation document (University of Hull, 2019), Hull Ltd identified six desired impact domains: cultural impact, place making, economic impact, social impact, and partnership development. Within each domain, Hull Ltd. identified causal pathways which they posited supported their overall impacts. For example, one identified pathway under its place making objective is written as follows: its capital program created an immediate output of enhanced public realm of the city center which would eventually lead to a longer-term outcome of visitors experiencing a higher quality city center which would eventually lead to the ultimate impact of having ‘people from within and outside Hull have an improved perception of the city’ (University of Hull, 2019, p.34). This was just one example of the several pathways which Hull Ltd. claimed to be creating this ultimate impact.

Necessitated by the weakening justification of public cultural funding, the cultural sector shifted its operation in an attempt to strengthen its case for investment. This led to the practice whereby cultural organizations used impact measurement to demonstrate their effect and legitimize claims to the public purse. While impact measurement was seen as a path towards legitimacy under NPM (demonstrated by both cultural organizations and an apparent uptake within DCMS itself) its usage was not without critique.

CRITIQUES OF IMPACT MEASUREMENT

As the cultural sector grew accustomed to using impact measurement to account for its value creation, so too grew a body of literature critical of how certain methods were deployed in practice. Critiques emerged that focused on measures of economic impact (Seaman, 2020; Throsby, 2003) as well as those of social impact (Soh, 2019; Belfiore & Bennett, 2009).

Counterfactual Arguments

Economic impact assessments used by the cultural sector have been criticized by economists to the point of being characterized as ‘a parody of economic analysis’ (Seaman, 2006). The main critique of cultural economic impact studies is their misuse of substitution effects within GVA techniques. Substitution effects are rooted in a longstanding understanding that ‘arts institutions vie for leisure-time dollars that might have been spent in the community even if they were not spent on the arts’ (Cwi & Lyall, 1977, p.4). Following this logic, money spent in the cultural sector may be redistributed from other local leisure activities and should not be counted as ‘new’ money that would not have been spent in the absence of cultural organizations. As one critic noted, ‘simply documenting the relative size of arts spending versus non-arts spending is a useful exercise in data gathering, but these effects should be distinguished carefully from claims related to the overall growth and health of the local economy’ (Seaman, 2000, p.276).

Miscalculating the impact of cultural organizations is not a benign practice, but has been evidenced to create significant waste of public monies. This is exemplified through the case of the Mountain Laurel Performing Arts Center (MLPAC) in the United States. The MLPAC was built in 2003 to be a regional venue for the performing arts. To convince government officials to contribute towards construction costs, investors commissioned an economic impact study of the potential venue (Sterngold, 2004). The study found that MLPAC would create 1,273 permanent new jobs, produce annual increases of \$34.8 million in household income, \$82.4 million in economic output, and \$668,000 in property taxes during its early years of operation (Zaricki 1993). The study accomplished its advocacy goal and \$36 million in government funds were allocated to the construction of the center. Seven months after opening, the center filed for insolvency due to mounting financial debt (Charlanza 2004). As researchers looked back to assess what went wrong with MLPAC, they found that the initial economic impact study used GVA techniques which ‘did not distinguish between net additions to regional spending and diversions of spending from substitute uses within the region’ (Sterngold, 2004, p.177). Had the initial economic impact study used more robust calculative techniques, policy officials may have been able to better allocate public funds.

In addition to misrepresenting the overall economic impact of their activity, the cultural sector has been accused of overclaiming the extent to which public cultural investment translates into economic return more than investments into other sectors. For example, Strom

& Cook (2004) note that ‘one cannot conclude that support for the arts, given particular economic goals such as the creation of jobs, is more desirable than other uses of public dollars’ (p.508). Characterizing this point beautifully, Peacock (2000) applies this critique to public bodies that fund cultural activity, saying:

Unless that body can conduct some franchising arrangement which induces competitive tenders that reveal costs, a classic principal–agent situation arises with asymmetric information. ...There is therefore no guarantee that the response to a unit subsidy would be a movement along a minimum marginal cost curve, for the opportunity exists to accumulate rents by disguising them as costs. Of course, any such suggestion would be strongly denied by the subsidised firm. The case is usually associated with situations where there is a proposal to limit costs, in which case it is invariably claimed that any reduction in subsidy will lower the quality of output (p.195).

Put another way, Peacock argues that public bodies are unable to verify that with more money invested, more results occur. Rather, current measurement is unable to argue how a £1 increase in funding translates into an ‘n’ increase in attributable impact.

Similar counterfactual critiques have been aimed at cultural contexts that account for their operations through social impact measurement (Soh, 2019; Belfiore & Bennett, 2009). As previously noted, social impact measurement accounts for the instrumental use of cultural activity on other social dimensions, such as addressing social inequalities and revitalizing our cities. What these measurements often overlook is how these instrumental goals might have been addressed through other, non-cultural means and to what extent those other interventions might prove more effective. For example, Guetzkow (2002) notes ‘instead of “what are the benefits of the arts,” the question becomes “what are the opportunity costs of using this money to fund the arts?” For example, are arts programs for at-risk youth more effective than the Boy Scouts or midnight basketball?’ (p.18). While the New Victory Theater example given above addresses this critique, our review of the literature found that this was by far the exception and not the rule.

Measurement Criticality

In addition to counterfactual critiques, other criticisms have emerged at how cultural contexts might fall within evaluative traps native to the social sciences during social impact measurement (Galloway, 2009; Guetzkow, 2002). This complaint is chiefly aimed at the

prevalence of selection bias within cultural activities, whereby cultural organizations may account for the causal creation of social benefits when, in fact, a correlation may already be in place to account for any found relationship. For example, scholarship has indicated that people who participate in cultural activities are healthier and happier than the general population (Costello 1998) while other research states that healthier people are more likely to volunteer in cultural programs (Thoits and Hewitt 2001).

Advocates of using social performance measures within cultural contexts posit that the use of more theoretical models of impact measurement might be able to produce more robust program assessments (Galloway, 2009). Critics argue that these tools might be capable of producing more accurate accounts, but are, at best, misused by managers who do not know how to properly use these complex measures. At worst, managers use these tools as a marketing strategy while pursuing their objectives in ways that they always have done (Belfiore, 2004). For example, the evaluation model produced for Hull Ltd. was written *after* the City of Culture intervention had already taken place (University of Hull, 2019).

Finally, a critique lodged at economic and social impact measures is that both often ignore the negative impacts that cultural organizations might create (Seaman, 2020; Belfiore, 2004). For example, events such as raves or rock concerts might produce negative impacts of noise pollution and delinquency which are often overlooked (a point responded to by Gazel's [1997] economic impact study of a Grateful Dead concert that assessed the local area's extra expenditures on security and crime related to the event). Moreover, the results of some social impact assessments directly conflict with each other, further casting doubt into how they were performed (see competing accounts of how local cultural organizations lead to lasting resident diversity [Stern, 1999] as well as gentrification [Zukin, 1989] in New York City).

While methods of impact measurement in cultural contexts have evolved over time, they have largely ignored mounting critiques. Indeed, our review uncovered recent examples of impact measurement where these deficits are in full display – such as Arts Council England's 2019 economic impact assessment using GVA techniques and Hull UK City of Culture's posthumous theory-based evaluation. Thus, while impact measurement is purported to justify public and private investment, current practices of impact measurement undermine these legitimacy claims amid deficits in criticality and counterfactual reasoning.

DISCUSSION

Our review uncovers inconsistencies between impact measurement justification and practice in cultural contexts. On one hand, impact measurement strengthens the justification for cultural investment by providing cursory evidence of produced outcomes (Galloway, 2009). On the other hand, impact measurement weakens this justification by failing to produce counterfactual accounts for how cultural spend achieves funding objectives (Belfiore, 2002). These inconsistencies can be theorized as a paradox of impact measurement in cultural contexts. We define this paradox as follows:

Impact measurement *both strengthens and weakens* the justification for public cultural investment.

This type of paradox is referred to as a ‘both/and’ paradox (Lewis, 2000) – whereby a stimulus (i.e., impact measurement) creates both an intuitive response (i.e., strengthening the case for cultural investment) and an unintuitive, opposite response (i.e., weakening the case for cultural investment) at the same time.

Our review uncovered that this paradox has its empirical roots in the antecedents, development, and evolution of impact measurement. Early justification for cultural investment stemmed from the belief that culture creates positive externalities. Critics slammed this justification as elitist and ‘income redistribution in reverse.’ At the same time, wider trends in ‘rational’ policy making led towards the use of impact measurement as a way to make funding decisions ‘more scientific’ and ‘less biased’ (Belfiore, 2004). Initial impact measurement practices often measured “reach” – the extent to which a cultural intervention was able to reach an intended outcome (e.g., number of audience members). However, we found that products of measurement, mainly in the form of textual evaluation reports, often fell apart under critical scrutiny and ignored the counterfactual conditions whereby public monies could have achieved desired results with greater efficiency through other means.

Our findings suggest that practices of impact measurement in cultural contexts align with NPM in rhetoric alone. While NPM seeks ‘the enhancement of accountability to customers for the quality of service via the creation of performance indicators’ (Power, 1997, p. 43), current practices of cultural impact measurement produce a myriad of performance indicators, but none of which speak to the efficiency or quality of service provided. For example, we discuss how cultural organizations often set performance targets for gross

audience attendance without accounting for how marginal units of investment effect audience size (Peacock, 2000). This practice does not align with tenets of efficiency associated with NPM but uses the language of performance targets to mask the continuation of justifying cultural investment on positive externalities.

Using NPM discourse to cover over the extension of historical justifications may also cast doubt on the extent to which the cultural sector can address social and instrumental uses of culture. For example, while a large swath of research explores how engaging with culture leads to health and wellbeing benefits (Taylor et al., 2015; Costello, 1998; Matarasso, 1997) none of these studies explored whether ‘culture for wellbeing’ initiatives are more efficient at producing health outcomes than other types of interventions. Advocates argue that even if culture for wellbeing initiatives are less efficient regarding health outcomes, they may still be able to create fringe benefits that are not created by their non-culture counterparts (Galloway, 2009). While this may be the case, unless cultural organizations can account for what specific additional benefit they may produce, instrumental uses are further justified by saying that culture is ‘good for society’ based on positive externalities.

Previous scholars have suggested that impact measurement might only be used within cultural organizations to produce ‘advocacy research’ (Belfiore & Bennett, 2010). Belfiore & Bennett (2010) suggests that ‘the methodologies that have been so far applied to the evaluation of the socio-economic impacts of the arts have been subjected to extensive criticism both by arts professionals... and academics, who have exposed the flaws, ideological bias and advocacy purposes lurking in many of them’ (p.122). Our study reveals a more dire situation: impact measurement may strengthen short-term advocacy efforts, but weaken its long-term justification. Since cultural investors seem satisfied with the sector’s current measurement activity (indeed, Arts Council England’s planned investment for 2018-2022 was £2.4 *billion* [Arts Council England, 2016a]), there is little incentive for the cultural sector to change the status quo. However, if future cultural investment becomes increasingly tied to achieving irrefutable impact, then cultural contexts might be managing themselves out of business.

Throughout our exploration we found several cases of cultural organizations that have misused performance measurement techniques; however, we also found unicorn cases where impact measurement techniques were used appropriately and further justified their investment (Maliekel et al., 2019). Thus, impact measurement techniques in-and-of themselves are not wholly inappropriate within cultural contexts, but the application of these techniques is widely

varied. This variance might be explained by the difficulty in capturing social impact. For example, it is easy to track tangible outputs (e.g., participant reach) but difficult to measure intangible impacts (e.g., participant wellbeing). Moreover, organizations that might have the capabilities to create social impact might not have the capabilities to measure that impact. Due to the technical variance of cultural contexts, the utilization of impact measurement as a decision-making instrument to rationalize investment should be heavily questioned.

CONCLUSION & RESEARCH AGENDA

While impact measurement in cultural contexts is seen as a way to justify investment, current practices both strengthen *and* weaken this justification. Previous scholarship focused on either the benefits (Soh, 2019) or deficits (Belfiore, 2002) of impact measurement. We are the first to move beyond this either/or distinction and conceptualize the found phenomena as a both/and paradox. By outlining this contradictory paradox, we hope to move beyond oversimplified and polarized notions of cost/benefit and recognize the complexity, diversity, and ambiguity of impact measurement in cultural contexts. Scholars have posited that paradoxes can either be *confronted*, *embraced*, or *coped with* (Lewis, 2000). We conclude our article by recommending future research aligned with these three approaches.

Confronting Impact Measurement. Future research could *confront* the identified paradox head-on by re-problematizing the underlying justification for impact measurement. Our review raises questions about the assumptions behind the practice of impact measurement and its ability to account for the means-ends relationships between interventions and their potential impact. In particular, our review questions whether cultural contexts can robustly measure their impact due to potential limits in their internal capabilities. Future scholars may wish to further problematize the relationship between impact measurement and investor justification to avoid the manipulation of information produced by measurement for advocacy purposes (Belfiore & Bennett, 2010) and to avoid the distraction of key resources away from their intended outcomes. Funders could also reconceptualize what counts as ‘data,’ allowing creativity in the production and presentation of impact evidence (e.g., Mennicken & Espeland, 2019). In total, instead of attempting to supplant traditional justifications for cultural investment, complexifying these arguments may provide a better template for policy makers

than the practice found through this review: outdated rationales guised by perfunctory measurement.

Embracing Impact Measurement. Escaping reinforcing cycles requires seemingly counterintuitive responses (Cameron & Quinn, 1988). To that end, cultural organizations and their scholars may endeavour to *embrace* this paradox and refine the tools of impact measurement in order to address the numerous critiques raised by our review. To do this, we recommend several methodologies that may be auspicious for these contexts.

Organizational Efficacy. Cultural organizations have been duly criticized for representing their performance in gross figures – using measures such as total number of audiences served or GVA techniques to calculate economic performance (Sterngold, 2004). Responding to these critiques, cultural organizations might benefit from defining program objectives ad hoc and outline how programmatic decisions lead to either more productive or less productive activities. However, since cultural organizations premise impact measurement on gaining investment, they may be loath to measure internal efficiency when current funders are more concerned with external outputs, even though such practices may delegitimize their claims to funds in the future.

Case Control Studies. Similar to how theory-based evaluation methodologies may account for the social performance of cultural organizations, case control methodologies may add additional rigor to the measurement process. Case control methodologies are native to the social sciences and compare exposure groups with control groups over the course of a studied intervention (Groenland & Dana, 2019). Addressing a key critique of current impact measurement practices in cultural organizations, case control studies are able to explore the relationship between correlation and causation. The use of case control methodologies is rarely applied within cultural organizations but provide key insights when conducted (e.g., Maliekel et al., 2019).

Explicit Counterfactuals. In addition to addressing facets of criticality, future impact measurement in cultural contexts might benefit from addressing the counterfactual deficit of current practices. Cultural organizations and their scholars should examine how cultural organizations achieve ‘instrumental’ outcomes as compared with other initiatives addressing similar aims. For example, research could be conducted to measure the health outcomes of ‘culture for wellbeing’ initiatives as compared with non-culture related initiatives. The results

of such studies could clarify for decision makers the appropriate allocation of fiscal resources when considering funding these instrumental uses. Cultural interventions may be reluctant to measure their performance in this way due to the fear that they could be defunded if they demonstrate less outcome creation than other types of interventions. Conversely, interventions could find that they are hyper-efficient at creating intended outcomes and could see an influx in revenue.

Coping With Impact Measurement. Openly and critically *cop*ing with the polarized contradictions within our found paradox may enable more dramatic changes to future understandings and behaviors. While confronting or embracing the found paradox may prove demanding, coping with the paradox may prove more realistic. One approach to coping with the paradox would be to reshape cultural funding relationships. Instead of requiring perfunctory measurement for all funding arrangements, funders could allow most recipients to receive funding with minimal reporting requirements. At the same time, funders could give additional funds to specific interventions to pilot innovative techniques of impact measurement in an attempt to develop nuanced impact studies that hold up to critical scrutiny.

Funding culture in this way would cope with the paradox of impact measurement in two ways. First, current perfunctory measurement weakens the case for cultural investment by failing to produce evidence that stands up to critical scrutiny. If perfunctory measurement is no longer required, the cultural sector may stem the flow of evidence it produces against itself. Second, investing additional funds into innovative measurement initiatives may produce novel understandings about cultural impact that may strengthen its investment justification. For example, funding a cultural initiative that seeks to conduct a case-control study to measure the effectiveness of its program. If funders would engage in such practices, they could strengthen their claim of being ‘scientific’ by publishing peer-reviewed articles based on this stronger evidence base. Future scholarship should examine how this and other models cope with the paradox of impact measurement in cultural contexts.

REFERENCES

- Alexander, V. D. (2018). Enterprise culture and the arts: neoliberal values and British art institutions. In *Art and the Challenge of Markets Volume 1* (pp. 67-93). Palgrave Macmillan, Cham.
- Arts and Humanities Research Council (2019). *Delivery Plan 2019*. Arts and Humanities Research Council: London.
- Arts Council England (2021). *Let's Create Delivery Plan 2021-2024*. Arts Council England: London.
- Arts Council England (2019). *Contribution of the arts and culture industry to the UK economy*. Arts Council England: London.
- Arts Council England (2016a). *Announcing our budget for our 2018 - 2022 investment period*. Arts Council England: London.
- Arts Council England (2016b). *2013-14 – NPO Annual Data Survey – Full Dataset*. Arts Council England: London.
- Barsdate, K. (1997). *Measuring your arts economy: Twelve questions and answers about economic impact studies*. Washington, DC: National Assembly of State Arts Agencies (NASAA).
- Belfiore, E. (2002). Art as a means of alleviating social exclusion: Does it really work? A critique of instrumental cultural policies and social impact studies in the UK. *International journal of cultural policy*, 8(1), 91-106.
- Belfiore, E. (2004). Auditing culture: The subsidised cultural sector in the New Public Management. *International journal of cultural policy*, 10(2), 183-202.
- Belfiore, E., & Bennett, O. (2007). Rethinking the social impacts of the arts. *International journal of cultural policy*, 13(2), 135-151.
- Belfiore, E., & Bennett, O. (2009). Researching the social impact of the arts: literature, fiction and the novel. *International Journal of Cultural Policy*, 15(1), 17-33.
- Belfiore, E., & Bennett, O. (2010). Beyond the “Toolkit Approach”: arts impact evaluation research and the realities of cultural policy-making. *Journal for cultural research*, 14(2), 121-142.
- Cameron, K. & Quinn, R. (1988). Organizational paradox and transformation. In R. E. Quinn & K. S. Cameron (Eds.), *Paradox and transformation: Toward a theory of change in organization and management*. (pp. 12-18). Cambridge, MA: Ballinger.
- Charlanza, B. (2004). *A Poconos' slot spot could save arts center*. Eastern Pocono Community News, July 30, 2004.
- Christensen, T., & Lægheid, P. (2016). *The Ashgate research companion to new public management*. Routledge.
- Cohen, R., Schaffer, W., & Davidson, B. (2003). Arts and economic prosperity: The economic impact of nonprofit arts organizations and their audiences. *The Journal of Arts Management, Law, and Society*, 33(1), 17-31.

- Costello, J. (1998). *The Economic and Social Impact of the Arts on Urban and Community Development*. Pittsburgh Press: University of Pittsburgh.
- Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American journal of Evaluation*, 32(2), 199-226.
- Cwi, D., & Lyall, K. (1977). *Economic impacts of arts and cultural institutions: a model for assessment and a case study in Baltimore: a report*. Washington D.C.: National Endowment for the Arts.
- Department for Culture, Media & Sport (2008). *Performance Indicator Guidance 2008*. DCMS, London.
- Edwards, A. (1996), *The British Museum: A Fundamental Review of the Museum's Operations*. London: British Museum.
- Fiske, E. B. (Ed.). (1999). *Champions of change: The impact of the arts on learning*. Arts Education Partnership.
- Galloway, S. (2009). Theory-based evaluation and the social impact of the arts. *Cultural trends*, 18(2), 125-148.
- Gazel, R. C., & Schwer, R. K. (1997). Beyond rock and roll: The economic impact of the Grateful Dead on a local economy. *Journal of Cultural Economics*, 21(1), 41-55.
- Guetzkow, J. (2002). How the arts impact communities: An introduction to the literature on arts impact studies. *Journal for Arts and Cultural Policy Studies*, 20, 1-26.
- Hadley, S., & Gray, C. (2017). Hyperinstrumentalism and cultural policy: means to an end or an end to meaning?. *Cultural Trends*, 26(2), 95-106.
- Hood, C. (1991). A public management for all seasons?. *Public administration*, 69(1), 3-19.
- Hutchinson, R. (1982). *The Politics of the Arts Council*. London: Sinclair Browne Ltd.
- Jenkins, S. (1995) *Accountable to None*. Hamish Hamilton, London.
- King, K., & Blaug, M. (1979). Does the Arts Council know what it is doing?. In M. Blaug (Ed.), *The economics of the arts*, (pp.101-125). Martin Robertson & Co. (Original work published 1976).
- Lewis, M. W. (2000). Exploring paradox: Toward a more comprehensive guide. *Academy of Management review*, 25(4), 760-776.
- Lynch, R. T., & Chosa, D. (1996). Group-oriented community-based expressive arts programming for individuals with disabilities: Participant satisfaction and perceptions of psychosocial impact. *Journal of Rehabilitation*, 62(3), 75.
- Madden, C. (2001). Using 'economic' impact studies in arts and cultural advocacy: A cautionary note. *Media International Australia Incorporating Culture and Policy*, 98(1), 161-178.
- Maliekel, L. B., Boddie, C. J., Wolf, D. P., & Holochwost, S. (2019). Theaters for Learning: The New Victory Theater SPARK Program. In *Education and Theatres* (pp. 195-207). Springer, Cham.
- Matarasso, F. (1997). Use or ornament?. *The social impact of participation in the arts*, 4(2).

- Mennicken, A., & Espeland, W. N. (2019). What's new with numbers? Sociological approaches to the study of quantification. *Annual Review of Sociology*, 45, 223-245.
- Myerscough, J. (1988). *The economic importance of the arts in Britain*. Policy Studies Institute.
- Netzer, D. (1978). *The subsidized muse: Public support for the arts in the United States*. Cambridge University Press.
- Peacock, A. (1976). Welfare economics and public subsidies to the arts. In M. Blaug (Ed.), *The economics of the arts* (pp.70-83). Martin Robertson & Co. (Original work published 1969).
- Peacock, A. (2000). Public financing of the arts in England. *Fiscal Studies*, 21(2), 171-205.
- Peaker, A., & Vincent, J. (1990). *Arts in prisons: a sense of achievement*. Loughborough University, Centre for Research in Social Policy.
- Power, M. (1997). *The Audit Society: Rituals of Verification*. Oxford University Press, Oxford.
- Quinn, R. (1997). Distance or intimacy?—The arm's length principle, the British government and the arts council of Great Britain. *International journal of cultural policy*, 4(1), 127-159.
- Rees-Mogg, W. (1984). *The Glory of the Garden: The Development of the Arts in England, a Strategy for a Decade*. Arts Council of Great Britain.
- Reid, D. (2009). An American Vision of Federal Arts Subsidies: Why and How the US Government Should Support Artistic Expression. *Yale JL & Human.*, 21, 361.
- Reith, J. (1924). *Broadcast over Britain*. London: Hodder and Stoughton.
- Scannell, P., & Cardiff, D. (1991). *A Social History of British Broadcasting: 1922-1939: Serving the Nation*. Basil Blackwell.
- Seaman, B. (2000). Arts impact studies: A fashionable excess. In *The politics of culture*, (eds.) Bradford G., Gary, M., and Wallach, G. New York: New York Press.
- Seaman, B. A. (2006). The relationship among regional economic impact models: contingent valuation versus economic impact in the case of cultural assets. *Andrew Young School of Policy Studies Research Paper Series*, 7(5).
- Seaman, B. A. (2020). Economic impact of the arts. In *Handbook of Cultural Economics, Third Edition*. Edward Elgar Publishing.
- Selwood, S. (2002). The politics of data collection: Gathering, analysing and using data about the subsidised cultural sector in England. *Cultural trends*, 12(47), 13-84.
- Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). MIT press.
- Soh, L. Y. (2019). Harnessing the Transformative Power of Arts and Culture for Social Impact. In *Arts Education and Cultural Diversity* (pp. 171-181). Springer, Singapore.
- Stern, M. (1999). *Is All the World Philadelphia? A multi-city study of arts and cultural organizations, diversity, and urban revitalization*. Philadelphia: University of Pennsylvania School of Social Work.

- Sterngold, A. H. (2004). Do economic impact studies misrepresent the benefits of arts and cultural organizations?. *The journal of arts management, law, and society*, 34(3), 166-187.
- Strom, E., & Cook, A. (2004). Old pictures in new frames: Issue definition and federal arts policy. *Review of Policy Research*, 21(4), 505-522.
- Taylor, P., Davies, L., Wells, P., Gilbertson, J., & Tayleur, W. (2015). A review of the social impacts of culture and sport. London: DCMS.
- Thoits, P., & Hewitt, L. (2001). Volunteer Work and Well-Being. *Journal of Health and Social Behavior* 42:115-131.
- Throsby, D. (2003). Determining the value of cultural goods: How much (or how little) does contingent valuation tell us?. *Journal of cultural economics*, 27(3-4), 275-285.
- Throsby, D. (2010). *The economics of cultural policy*. Cambridge University Press.
- Throsby, D. (2017). In defence of measurement: a comment on Phiddian et al. *Cultural Trends*, 26(4), 314-317.
- Throsby, D. & Withers, A. (1979). *The Economics of the Performing Arts*. London: Edward Arnold.
- University of Hull. (2019). *Cultural Transformations: The impacts of Hull UK city of culture 2017*. Hull: University of Hull Culture, Place, and Policy Institute.
- Upchurch, A. (2004). John Maynard Keynes, the Bloomsbury group and the origins of the arts council movement. *International Journal of Cultural Policy*, 10(2), 203-217.
- Upchurch, A. R. (2016). *The origins of the Arts Council movement: philanthropy and policy*. Springer.
- Zaricki, S. (1993). *Economic impact of the Mountain Laurel Center for the Performing Arts in Pike County*. Pittston: Northeastern Pennsylvania Alliance.
- Zukin, S. (1989). *Loft Living: Culture and Capital in Urban Change*. New Brunswick: Rutgers University Press.