



Research Repository

The emergence of the thoughtless university: an Arendtian perspective on the Italian third mission metric regimes

Accepted for publication in Accounting, Auditing & Accountability Journal

Research Repository link: <https://repository.essex.ac.uk/43625/>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the published version if you wish to cite this paper.

<https://doi.org/10.1108/AAAJ-02-2026-8924>

The Emergence of the Thoughtless University: An Arendtian Perspective on the Italian Third Mission Metric Regimes

Abstract

Purpose. This paper aims to understand how academics experience the metrification of Third Mission activities and how these evaluative processes reshape the conditions for labour, work, and action within the contemporary university.

Design/methodology/approach. The study adopts an autoethnographic reflection and semi-structured interviews with academics, university leaders, Third Mission delegates, and early-career researchers operating within the Italian higher education system.

Findings. The findings reveal that Third Mission evaluation generates growing procedural burdens, uncertainty regarding evaluative criteria, inequities in recognition and workload allocation, and a widening gap between institutional narratives of societal value and academics' everyday practices. Metric-driven evaluation reorients engagement activities toward procedural compliance, privileging what is reportable over what is relational, ethical, and transformative. These findings suggest the need for evaluation systems that recognise the qualitative, relational, and ethical dimensions of societal engagement and that preserve academic agency, plurality, and responsibility.

Originality/value. The paper contributes to interdisciplinary accounting research by extending debates on metrification and accountability to universities' societal engagement. By mobilising Hannah Arendt's distinction between labour, work, and action, it conceptualises the Third Mission evaluation as a process that risks transforming academics into labourers of metric compliance, eroding spaces for thinking, meaning, and genuine public action. The study offers a novel theoretical lens for understanding how performance regimes reshape academic subjectivities and public value.

Keywords

Third Mission; performance measurement; metrification; societal engagement; higher education; accountability; Hannah Arendt; academics

1. Introduction

Universities across Europe have undergone profound transformations under increasingly pervasive regimes of performance measurement, evaluation, and audit. A substantial body of research has shown how these shifts reshape institutional logics, academic identities, and the relationship between universities and society (Martin-Sardesai et al., 2020; Parker, 2024; ter Bogt & Scapens, 2012). Scholars have noted that the metrification of academic work risks narrowing the meaning of value, privileging what is measurable over what is meaningful (Michelon, 2021; Gendron & Rodrigue, 2021; Carnegie, 2022, 2023). This dynamic is associated with unintended consequences such as identity tensions, metric distortion, strategic compliance, and the erosion of academic freedom (Kallio et al., 2017a; Argento & van Helden, 2023; Gebreiter, 2022). While these critiques have traditionally focused on research assessment, teaching evaluation, and funding mechanisms, a parallel transformation is unfolding around universities' so-called Third Mission activities. Third mission activities, encompassing public engagement, technology transfer, cultural dissemination, lifelong learning, and social innovation, are becoming a central element in higher education activities, as well

as performance regimes (Sánchez-Barrioluengo, 2014). This paper sets out to explore the implications of the expansion of metric governance into the Third Mission.

As universities and academics, especially in specific contexts such as Italy, are requested to report about their Third mission activities, performance measurement extends into another domain of academic work. This development has the potential to have direct implications for academics' everyday work, affecting how time is allocated, how activities are prioritised, and what becomes recognised as legitimate contribution within institutions. As Third Mission activities become more visible within evaluation systems, they are likely to increasingly affect how academics allocate time, document their work, and decide which forms of engagement are worth pursuing. They also potentially influence what becomes recognisable as legitimate contribution within departments and institutions, often privileging activities that are easily reportable, auditable, or strategically aligned while rendering more informal, relational, or long-term forms of engagement less visible. In this sense, the metrification of the Third Mission does not simply create an additional administrative task; it also contributes to reshaping academic priorities, workload distribution, and the practical meaning of engagement itself. Nevertheless, and despite the important insights provided by the studies mentioned above, little attention has been paid to how the metrification of the Third Mission is experienced by the academics responsible for enacting it in their everyday work. We know relatively little about how the evaluative infrastructures underpinning the Third Mission reshape academics' understandings of societal engagement, reorganise their temporal and practical priorities, and redefine the conditions under which meaningful engagement can take place.

The objective of this study is to understand how academics experience the metrification of Third Mission activities and how these evaluative processes reshape the conditions for labour, work, and action within the contemporary university. Addressing this objective requires moving beyond the formal architecture of performance measurement to examine how evaluative regimes are lived, interpreted, and negotiated in everyday academic practice. To this end, the study addresses the following research question:

How do academics experience the metrification of Third Mission activities, and how does this process reshape the conditions for labour, work, and action within the university?

The paper adopts an Arendtian perspective, which enables us to conceptualise metrification not merely as a managerial or technical development but as a process that transforms the conditions of thinking, judgment, and action within bureaucratically organised settings (see also Holt, 2020). Drawing on Arendt's distinction between labour, work, and action, we examine how metric-driven evaluation reconfigures academics' relationship with societal engagement, progressively privileging repetitive compliance activities and the production of evaluative artefacts while constraining the possibilities for plurality, judgment, and meaningful public action.

Italy provides a particularly revealing setting in which to investigate these dynamics. Here, the institutionalisation of the Third Mission has been accompanied by the progressive development of formal evaluation systems that are oriented less towards assessing the societal effects of university engagement than towards documenting, classifying, and governing a wide range of university–society interactions. To avoid conceptual ambiguity, we distinguish between Third Mission activities (TMA) - the substantive practices of societal engagement undertaken by academics and universities - and Third Mission metric regimes (TMMR) - the evaluative architectures, indicators, reporting systems, and accountability arrangements through which those activities are rendered visible, comparable, and governable. The paper therefore investigates how TMA become progressively mediated through TMMR and how academics experience this transformation in their everyday work.

Methodologically, the study combines autoethnographic reflection with fourteen semi-structured interviews involving academics, university leaders, Third Mission delegates, and early-career researchers embedded within the Italian higher education system (Alvesson, 2003; Haynes, 2011). This combination of methods allows us to capture both the institutional dynamics and the lived experiences of metrification, showing how Third Mission evaluation reorients academic work towards recurrent cycles of reporting and compliance, channels engagement into the production of evaluative artefacts, and narrows the institutional conditions under which action, understood in Arendtian terms as plural, dialogic, and publicly meaningful engagement, can flourish.

The findings go beyond existing critiques of performance measurement by showing how metrification reshapes academic practice in the underexplored domain of societal engagement. While prior studies have focused on technical limitations, incentive distortions, or symbolic compliance, we show how evaluative regimes reconfigure the conditions under which academics can engage meaningfully with society. The fundamental misalignment is that, while the Third Mission depends on meaningful impact, current evaluative regimes privilege what is measurable, reportable, and auditable. Consequently, societal engagement is increasingly transformed into repetitive labour oriented towards metric compliance, constraining the realm of action and risking foreclosing the Third Mission's own promise. Along these lines, we suggest that the metrification of the Third Mission reinforces a condition of thoughtlessness in academia, whereby governance arrangements privilege procedural alignment over reflective judgment about what engagement is for and whom it serves.

By reframing Third Mission evaluation through Arendt's thought, the paper offers a more focused contribution than traditional critiques of performance measurement. Rather than rejecting evaluation *per se*, it shows that when societal engagement is reduced primarily to measurable outputs, the institutional conditions for judgment, plurality, and meaningful public action may be weakened. At the same time, the findings suggest that evaluation systems also hold the potential to support richer forms of accountability, provided they recognise the qualitative, relational, and deliberative dimensions of engagement alongside quantitative indicators. The study thus extends debates on accountability to universities' societal role, highlighting how evaluation systems reshape academic identities, institutional purposes, and notions of public value. At its core is the concern that current regimes may foster forms of "thoughtlessness," understood as a separation between thinking and action. Yet, from an Arendtian perspective, this diagnosis also points towards a constructive alternative. Reclaiming the Third Mission requires designing evaluation arrangements that preserve spaces for reflection, plurality, and action, enabling academics to engage with society as responsible and reflective actors rather than merely as producers of indicators. In this sense, the paper does not advocate abandoning evaluation but reimagining it as an infrastructure that can cultivate judgment, sustain democratic dialogue, and reinforce the university's public mission.

The remainder of this paper is structured as follows. Section 2 assesses prior studies. Section 3 illustrates the theoretical framework. Section 4 describes the methodological approach, while section 5 is devoted to the study context. The findings are presented in section 6, while section 7 provides a theoretically-framed discussion. Section 8 provides some concluding remarks and implications.

2. The evolution of performance measurement in Universities/Higher Education towards the Third Mission: a literature review

Scholarly interest in performance management in higher education has grown substantially over the past two decades, driven by widespread reforms that have embedded evaluative mechanisms, metric-based accountability, and managerial logics into university governance systems. Previous analyses noted the emergence of new institutional logics centred on efficiency, quantification, and competition, challenging traditional ideas of academic autonomy and collegial self-governance (ter Bogt &

Scapens, 2012; Martin-Sardesai et al., 2020). Studies have since documented how performance measurement reshapes academic behaviour, influences research agendas, and introduces forms of control that often conflict with the intrinsic motivations associated with scholarly work (Parker, 2024; Gendron & Rodrigue, 2021). Metrification, in particular, has attracted critical attention for its tendency to reduce complex academic activities into simplified indicators that privilege what is measurable over what is meaningful (Michelon, 2021; Gebreiter, 2022; Carnegie 2022, 2023)). Several studies highlight the tensions and contradictions embedded in metric-driven governance systems, where institutional priorities are increasingly shaped by externally defined indicators and competitive positioning (Carnegie, 2022; Parker et al., 2024). This dynamic raises concerns about unintended outcomes such as metric gaming, symbolic compliance, role overload, and the erosion of academic freedom (Aboubichr & Conway, 2023; Argento & van Helden, 2023; Kallio et al., 2017a). In this respect, critiques of global university rankings similarly underline how metric regimes tend to privilege visibility, comparability, and symbolic performance over substantive contributions, reinforcing a disconnection between what is measured at the system level and what is experienced and enacted within academic practice (Carnegie, 2022). These contributions point to what has been described as “macro-micro contradiction to public university management” (Carnegie, 2022, *Global university rankings*, p. 77), whereby system-level performance imperatives conflict with the everyday realities, values, and practices of academic work.

Within this broader transformation, increasing attention has been devoted to universities' societal engagement, conceptualised as impact, Third Mission, knowledge valorisation, or public engagement. Although increasingly recognised as a strategic means through which universities demonstrate societal relevance and contribute to economic and social development (Sánchez-Barrioluengo, 2014), this domain remains less developed than research and teaching in the performance measurement literature. Consequently, governments and university leaders have progressively integrated Third Mission evaluation into quality assurance systems, funding models, and institutional strategies (Benneworth & Jongbloed, 2010).

The literature, however, raises several concerns regarding these developments. Scholars note that such activities are often characterized by relational, contextual, and long-term engagements that do not easily align with standardized indicators or metric-driven assessment frameworks (Andersen et al., 2016). Efforts to measure societal impact risk privileging activities that produce visible, short-term, or quantifiable outputs, while undervaluing forms of engagement that rely on trust-building, co-creation, or cultural and civic significance (Steccolini, 2019). Recent work in public sector accounting emphasises that value creation in public organisations is inherently plural, contested, and embedded in specific contexts, and therefore cannot be adequately captured through narrow performance indicators (Bracci et al., 2021; Sicilia et al., 2025). This misalignment may lead academics to prioritize activities that “count” in evaluative systems over those that they believe hold substantive societal value. This prioritisation is not necessarily driven by intrinsic belief in the metrics, but by structural incentives and risks. Activities that align with evaluative criteria influence resource allocation, institutional visibility, and, indirectly, career trajectories. Conversely, activities that do not “count” risk remaining invisible, unrewarded, or even questioned. Under such conditions, academics rationally adjust their engagement strategies to align with institutional signals. Furthermore, as performance-based accountability expands to incorporate the so-called Third Mission, academics face additional responsibilities on top of research, teaching, grant acquisition, and administrative commitments. Several studies highlight how these accumulating demands risk producing workload inflation, inequities in recognition, and tensions between institutional expectations and professional identities (Manes Rossi et al., 2022; Spanò et al., 2024; Argento & van Helden, 2020).

A recurring theme is the role of internal stakeholders, particularly academics, as both agents and subjects of Third Mission processes. While policy narratives emphasise societal engagement, the

practical implementation of Third Mission strategies often unfolds through internal managerial systems that define what “impact” is, how it should be documented, and which activities are institutionally rewarded. Research on stakeholder engagement in public organizations underscores the importance of dialogic, participatory approaches capable of recognizing multiple perspectives and forms of value (Morsing & Schultz, 2006). Yet, in the context of universities, empirical studies frequently report limited involvement of academics in shaping Third Mission evaluation frameworks, contributing to disconnection, frustration, or symbolic compliance (Spanò et al., 2022). In the Italian context, academics, department delegates, and university-level coordinators play a crucial role in translating policy expectations into everyday practices, deciding what activities are reported, how they are framed, and which forms of engagement become institutionally visible. Focusing on such internal stakeholders allows us to examine the micro-level processes through which evaluative regimes shape societal engagement from within, before it reaches external publics.

In sum, the literature points to a growing tension between the aspirations associated with societal engagement and the realities of metric-driven performance regimes. While Third Mission policies aim to legitimize universities’ contributions to society, the mechanisms used to assess these contributions risk reducing their richness and complexity. More broadly, this debate can be situated within longer-standing critiques of New Public Management (NPM) and its enduring influence on public sector governance. While NPM-inspired reforms have been widely associated with the expansion of performance measurement and audit practices since the 1980s, critical scholarship has repeatedly questioned their capacity to account for the complexity and contextual nature of public organisations, including universities. Hopwood’s (1990) work emphasises that accounting systems are not neutral tools but are deeply embedded in the organisational and social contexts in which they operate. From this perspective, the metrification of the Third Mission can be seen as part of a broader tendency to apply standardised evaluative frameworks to activities that are inherently relational, context-dependent, and value-laden (Loi & Di Guardo, 2015). This reinforces concerns that contemporary performance regimes, including those governing societal engagement, risk reproducing long-standing limitations of NPM by privileging formal accountability and measurement over substantive understanding and responsibility.

This tension highlights the need for theoretical lenses capable of interrogating not only the managerial and technical aspects of Third Mission evaluation but also its implications for academic identity, independence, agency, and meaning. The expansion of evaluative regimes does not simply introduce new tools; it alters the relationship between means and ends within academic work. Activities that were previously pursued for their intrinsic societal value increasingly become oriented toward satisfying external evaluative requirements. In this process, the demonstration of engagement risks becoming more important than engagement itself. This inversion - where procedures and indicators take precedence over meaning and judgment - cannot be fully captured through managerial or technical critiques alone. It calls for a conceptual framework capable of interrogating how evaluative systems reshape the conditions under which academics can think, judge, and act. Hannah Arendt’s distinctions between labour, work, and action offer precisely such a lens and is adopted in this investigation.

3. Rethinking Metrification through Arendt: Labour, Work, and Action

This paper adopts an Arendtian perspective to examine the metrification of the Italian Third Mission activities, as it allows us to interrogate how metric regimes reshape the conditions under which academics think, judge, and act. It foregrounds the question of whether and how contemporary evaluative arrangements contribute to what Arendt (1958) describes as “thoughtlessness”, understood not as an individual failing, but as an institutional condition in which procedural conformity displaces reflective judgment.

Hannah Arendt's political and philosophical writings provide a rich vocabulary for interrogating how contemporary universities structure, constrain, and make possible particular forms of academic activity. Although Arendt did not write about universities or performance measurement directly, her reflections on the conditions that support human plurality, judgment, and responsibility resonate strongly with the dynamics produced by metric-driven governance. Her work enables an analysis that moves beyond managerial or technical critiques, allowing us to examine how universities' evaluative infrastructures reshape the very texture of academic life and the possibilities for meaningful societal engagement. In this regard, Holt (2020) discusses Arendt's argument by showing how modern educational institutions cultivate "atmospheric thoughtlessness" through three interlocking tendencies: groupthink, proceduralism, and vocationalism. These tendencies emerge when institutions prioritise alignment, standardisation, and performativity over critical plurality. They are not merely pedagogical shortcomings but indicators of a deeper crisis in which evaluative routines corrode the capacity to think and judge. As Holt (2020) notes, modern educational institutions increasingly orient activities toward "questions of knowledge" - what can be measured, taught, certified, and controlled - rather than "questions of meaning" that require judgment, plurality, and critical reflection. This shift parallels Arendt's movement from action toward work and labour: evaluative knowledge systems privilege certainty and procedural compliance, thereby narrowing the conditions under which action and meaning can emerge. Therefore, we build upon Holt's suggestion and explain in detail central concepts in Arendtian works, to frame our understanding of how the metrification of Third Mission activities is lived, interpreted, negotiated, and resisted by academics.

A central contribution of Arendt's thought is the tripartite distinction between *labour*, *work*, and *action* articulated in *The Human Condition* (Arendt, 1958). This is not a classificatory scheme but a phenomenological account of how different forms of activity relate to necessity, durability, and meaning. *Labour*, for Arendt, designates cyclical and repetitive activity whose products are quickly consumed and must be continually reproduced. It is tied to necessity rather than freedom and does not contribute to the creation of a shared world. In contemporary universities, this notion helps articulate why academics frequently describe performance-related tasks as exhausting and devoid of lasting value. Reporting, uploading, documenting, updating indicators, and aligning activities with shifting evaluative categories exhibit precisely the qualities of labour: they must be performed endlessly, they are demanded by institutional necessity, and they seldom accumulate into anything recognised as meaningful. Critical accounting scholarship has highlighted similar patterns, noting the proliferation of symbolic compliance, metric gaming, and overload within audit regimes (e.g., Kallio et al., 2017b; Gebreiter, 2022; Spanò et al., 2024). Arendt's conceptualisation explains the sense of depletion conveyed in such studies: these practices are experienced as labour because they absorb energy without contributing durable value to the world academics inhabit.

If labour captures repetition and necessity, Arendt's notion of *work* illuminates the construction of the institutional artefacts that give stability and structure to human activity. Work produces durable objects, tools, infrastructures, categories, rules, that organise the world and define what can appear within it. In the context of the Third Mission, the evaluative infrastructures that universities build, such as templates, portals, indicators, quality assurance mechanisms, strategic plans, function precisely as such artefacts. They stabilise definitions of what counts as engagement and make particular forms of value visible while marginalising others. Prior research on the institutionalisation of performance measurement shows how these artefacts formalise complex practices and reshape organisational and personal priorities (Benneworth & Jongbloed, 2010). Through an Arendtian lens, they can be understood not merely as administrative devices but as elements of a fabricated world that academics must inhabit. Their durability and prescriptive force explain why academics often find themselves negotiating, translating, or bending their activities to fit predetermined categories. The possible resulting misalignments are consequences of world-building processes that privilege visibility, quantifiability, and comparability over relational, contextual, or emergent forms of value.

Arendt's third mode of activity - *action* - foregrounds the relational, dialogic, and unpredictable character of human engagement. To act is to introduce into the public realms something which, though intangible, is perfectly real and has consequences of its own. Every action is a new (unexpected) beginning that foregrounds "the freedom to call something into being which did not exist before". Action takes place in the presence of others; it is through action that individuals disclose who they are and participate in the creation of a shared world. In the university context, action captures the forms of intellectual and societal engagement that rely on plurality, trust, and co-creation rather than on the mechanical production of outputs. Much of what the Third Mission aspires to, that is collaboration with communities, dialogic modes of knowledge production, civic and cultural engagement, belongs to this sphere. The stakeholder and dialogic accountability literature similarly emphasises the importance of responsiveness, plurality, and relational value (Morsing & Schultz, 2006; Steccolini, 2019). Yet the same scholarship has shown how these practices are often overshadowed by managerial demands for demonstration, documentation, and standardisation. Recent advances in dialogic accounting further emphasise how accountability practices must recognise plurality, dissent, and relational forms of value creation, rather than being reduced to procedural or report-driven exercises (Tanima et al., 2024). Arendt's framing helps to explain this tension: the very qualities that make action meaningful - its unpredictability and plurality - sit uneasily with the desire for control and ordering that animates contemporary evaluative regimes.

Alongside these distinctions, Arendt's reflections on thinking and the dangers of thoughtlessness deepen the interpretive power of her framework. In *The Life of the Mind* (Arendt, 1978), she describes *thinking* as an internal dialogue that enables individuals to question assumptions, resist conformity, and form judgments about what is right. *Thoughtlessness*, by contrast, emerges when institutional routines encourage mechanical compliance. Her analysis of the "banality of evil" in *Eichmann in Jerusalem* (Arendt, 1963) is an extreme illustration of how bureaucratic environments can erode judgment by rewarding obedience over reflection. The relevance of this insight for universities lies not in moral equivalence but in institutional dynamics: where evaluative infrastructures proliferate, people increasingly confront pressures to follow procedures, satisfy indicators, and demonstrate compliance, sometimes at the expense of their own judgment. As critical accounting studies have observed similar dynamics of ritualism, proceduralism, and decoupling (e.g., Deem & Brehony, 2005; Manes Rossi et al., 2022), Arendt's distinction allows us to clarify how evaluative practices have the potential of displacing reflection, meaningful engagement and action.

Taken together, these concepts form a coherent interpretive framework for the present study. They provide a vocabulary for analysing not only what academics do in response to metric-driven expectations but also how they experience and make sense of these demands. The distinction between labour, work, and action sensitises the analysis to the different temporalities and meanings attached to academic activities, helping to reveal how metrification redistributes attention and energy across these spheres. The notions of thinking and thoughtlessness illuminate the conditions under which people can exercise judgment or are subtly encouraged to suspend it. Crucially, this conceptual apparatus enables an exploration of the lived realities of metrification that moves beyond managerial or technical critique. It reveals how evaluative artefacts shape the world in which people act, and how repetitive tasks have the potential to drain the capacity for meaningful engagement.

4. Methodological Approach

This study employs a qualitative, interpretive methodology designed to capture the lived experiences of academics as they navigate expanding evaluative demands associated with universities' Third Mission. The methodology relies on situated understanding, reflexivity, and plurality, elements that resonate with Arendt's conceptualisation of work, labour, and action. Methodologically, the inquiry combines autoethnographic reflection (Alvesson, 2003; Haynes, 2011) with semi-structured

interviews conducted with actors occupying diverse positions within Italian universities. This blended approach illuminates both the structural forces that shape Third Mission practices and the everyday negotiations through which individuals interpret and respond to them.

Autoethnography forms an integral part of the research design. While still relatively rare in accounting research, autoethnography has been employed to examine identity, professional rituals, and emotional labour (Haynes, 2013; Sadler & Wessels, 2019; Dar et al., 2025; Drujon d'Astros et al., 2025). These studies demonstrate the method's capacity to illuminate how institutional structures are experienced and negotiated from within. The present study extends this tradition by applying autoethnographic reflection to the context of performance measurement and societal engagement in universities. The researchers' long-term involvement in Italian higher education, as faculty members, committee participants, Third Mission delegates, and actors within evaluative processes, provides a privileged vantage point from which to observe the micro-politics of institutional change and the lived tensions surrounding societal engagement. Following Alvesson (2003), the approach treats the researchers' own experiences as meaningful empirical material, not as bias needing to be minimised but as a source of insight into the dynamics of organisational life.

Moreover, the authors' team is composed of authors who daily experience TMA roles and others with no active involvement, which was an equal balance (50/50) to avoid bias and excessive personalisation. Field notes and internal documents (such as departmental strategic plans, institutional emails, and reporting templates) were used to trace moments of uncertainty, dissonance, negotiation, and resistance. For example, one researcher documented repeated confusion during annual TM updates about "what counts" as engagement, capturing a sense of ambiguity later echoed in interviews. These autoethnographic materials were not treated as anecdotal but as part of a systematic reflexive practice that complements the interview data. Following Haynes (2017) and Zou (2021), autoethnography is not used here as confessional narrative but as an analytically structured reflexivity, in which personal experience becomes a lens for examining institutional processes. The study mobilises two interrelated layers of reflexivity. First, the researchers' own autoethnographic reflections draw on lived experience within Italian universities, documenting moments of ambiguity, negotiation, and tension as they unfolded in real time. These reflections were recorded contemporaneously in field notes and revisited analytically during coding. Second, interviewees themselves engaged in reflective sensemaking, articulating their own interpretations of institutional change. The analysis therefore captures both institutional dynamics, and reflexive processes embedded within academic culture.

The empirical material consists of fourteen semi-structured interviews conducted with academics, university leaders, Third Mission delegates, and early-career researchers, all of whom were directly or indirectly involved in Third Mission activities and exposed to related evaluation processes. Semi-structured interviews were adopted to explore participants' experiences and interpretations while allowing sufficient flexibility to probe emerging themes and capture the complexity of Third Mission practices. The number of interviews reflects a strategy of analytical depth rather than breadth. Data collection proceeded until thematic saturation was reached, that is, until additional interviews did not generate substantially new insights in relation to the research question. All of the participants belong to one of a total of four universities within the same territorial context, a large metropolitan area in Southern Italy that represents a paradigmatic case for third mission activities and their measurement, given its dense academic ecosystem and socio-economically complex environment, which allows for a consistent and context-sensitive analysis of knowledge transfer, innovation, and societal engagement impacts. This intends to ensure comparability in terms of regulatory and policy conditions, while allowing for variation in organisational characteristics such as size, disciplinary composition, and internal governance structures. This design enables the analysis to capture how similar evaluative frameworks are experienced across different institutional settings, without

introducing excessive variability related to broader contextual differences. This approach is aligned with interdisciplinary traditions in accounting research, where the richness of contextualised understanding is prioritised over sample size.

Participants were selected through purposive sampling, based on their engagement with Third Mission activities and their familiarity with institutional evaluation practices. The aim was not to achieve representativeness in a statistical sense, but to capture a diversity of perspectives across roles, career stages, and organisational positions. Importantly, interviewees were drawn from a range of disciplinary backgrounds, rather than from a single field such as accounting, reflecting the cross-cutting nature of Third Mission activities within universities. The combination of different roles (e.g., faculty members, delegates, and leaders) and perspectives enabled the identification of recurring patterns as well as tensions and divergences across the data. The sample included a balanced gender distribution (seven women and seven men), although the analysis does not aim to draw gender-based comparisons which were not addressed herein.

Interview durations ranged from approximately 20 to 90 minutes, allowing for both focused and more in-depth accounts depending on participants' roles and availability. Shorter interviews were typically associated with participants with more limited availability or more focused roles, while longer interviews allowed for more in-depth exploration of experiences among those more directly involved in Third Mission coordination and governance. Attention was also paid to ensuring that perspectives from participants with shorter interviews were given equal consideration, especially when they provided contrasting or disconfirming views. For participants holding multiple roles (e.g., academic and administrative responsibilities), attention was paid to how these positional overlaps shaped their perspectives. These actors are not neutral observers: they possess situated knowledge shaped by their embeddedness within evaluative regimes, broadly as members of the Italian academic community and in direct/indirect exposure to TMA and its evolving metrics.

On this ground, the data collected, analysed and interpreted, although rooted in the Italian context have potential to speaking to academic communities exposed to multiple evaluative regimes across different countries, as well as to public sector institutions in general affected by similar increasing pressures for a metric-based accountability. Given the researchers' own involvement in the Italian higher education system, particular attention was paid to reflexivity throughout the research process. While this positionality provided valuable contextual insight, it also required careful consideration to avoid shaping participants' responses or imposing preconceived interpretations. To mitigate this risk, interviews were conducted using open-ended questions designed to allow participants to articulate their own perspectives, including critical or divergent views. The analysis further relied on systematic comparison across interviews and attention to inconsistencies and tensions in the data, rather than privileging confirmatory evidence. Moreover, while documentary materials informed the interpretation and contextualisation of findings, also assuring triangulation among the various sources, the empirical analysis presented in the paper primarily draws on interview data and autoethnographic notes, from which illustrative quotes are reported.

Table 1 provides an overview of the participants' characteristics, including gender, role, and institutional affiliation. While the interviews included both male and female participants and a range of seniority levels, no systematic differences in perspectives emerged along these dimensions. Rather, participants' accounts were broadly consistent across gender and career stages, with variations more closely related to their involvement in Third Mission activities and exposure to evaluation processes. Similarly, although participants were drawn from universities with different organisational profiles, the institutions involved are comprehensive universities combining both research and teaching missions, rather than strictly research-intensive or teaching-oriented institutions. As such, the analysis

does not focus on institutional typologies *per se*, but on shared experiences of metrification within a common evaluative framework.

Table 1. Interviews

Interviewee	No.	Gender	Encoding	Duration minutes
Dean of the department	3	Male	D1	45
		Male	D2	25
		Female	D3	35
University professors involved in the Third Mission	2	Male	P1	30
		Female	P2	45
University early career researchers involved in the Third Mission	3	Female	R1	20
		Female	R2	30
		Male	R3	45
University delegate for the Third Mission	3	Female	UD1	60
		Male	UD2	30
		Male	UD3	75
Department delegate for the Third Mission (Academic)	3	Female	DD1	90
		Female	DD2	70
		Male	DD3	55
	14			655

Interviews were audio-recorded, transcribed, and anonymised. The analysis was characterised by an iterative movement between empirical material and theoretical concepts (Alvesson & Kärreman, 2007), developing through a continuous dialogue between data and theory. Initial themes emerging from the interviews and autoethnographic material were first identified inductively and then interpreted through the lens of Arendt's concepts of labour, work, and action, while these theoretical concepts were progressively refined in light of empirical observations. This iterative process ensured that Arendt's framework did not function as a predetermined coding template. Rather, the concepts of labour, work, action, judgment, and thoughtlessness were used as sensitising concepts, that is, as analytical lenses that directed attention to particular dimensions of participants' experiences without requiring the data to fit predefined theoretical categories. They informed the interpretation of emerging patterns, while allowing unexpected themes and tensions to remain visible and shape the subsequent analysis. Throughout the research process, ethical principles were carefully observed. Participants were informed of the aims of the study, reassured of confidentiality, and given opportunities to clarify or retract statements. Identifying details were removed from transcripts and autoethnographic notes. The researchers' insider status required ethical vigilance, including restricting autoethnographic material that could inadvertently reveal institutional identities.

The coding process was conducted manually, allowing for close engagement with the data and continuous refinement of categories. Initial open coding identified recurring themes in participants' accounts, which were then progressively organised and interpreted in relation to the theoretical framework. This approach prioritised depth of interpretation over procedural standardisation. The coding strategy was organised around two analytical levels. Level 1 codes correspond to theoretically

informed macro-categories derived from Arendt’s conceptual distinctions and were used as sensitising concepts. Level 2 codes capture recurrent empirical patterns emerging inductively from interview data and autoethnographic material. The table below reports the consolidated coding structure used in the analysis.

For example, the code “labour-like repetitive tasks” emerged from participants’ descriptions of recurring reporting and compliance activities, such as repeatedly completing similar forms and documentation. As one participant noted, “Every year we redo the same forms. It feels like cleaning the same room over and over again” (DD1). This empirical pattern was subsequently interpreted through the lens of Arendt’s concept of labour, understood as cyclical and repetitive activity oriented toward procedural reproduction rather than lasting outcomes. In a similar way, participants’ accounts of engaging with evaluation templates such as those associated with VQR or AVA3 highlighted how these instruments impose predefined categories and formats through which societal engagement must be articulated. These practices were interpreted as “work-like,” in the sense that they involve the production of relatively durable evaluative artefacts that shape how engagement is represented and understood within the institution. Alongside these patterns, several themes emerged inductively from the data, including “wanting to do more than the system values,” “relational engagement as invisible,” “fear of misalignment,” “professional identity stretched toward compliance,” and “frustration with definitional ambiguity.” These themes capture participants’ attempts to reconcile their own understanding of meaningful engagement with the constraints imposed by evaluative systems.

To illustrate, a typical coding sequence involved first identifying a raw excerpt, followed by inductive coding, which involved finding connections with Arendtian concepts and adding notes. An example of such a sequence is as follows: (i) Raw excerpt: “I spend more time explaining why my public lecture does not fit their categories than actually preparing it.”; (ii) Inductive code: “category mismatch frustration”; (iii) Linked Arendtian concept: “work imposing constraints on action”; (iv) theoretical note: “Participant describes institutional structures narrowing meaning, aligns with Arendt’s idea of fabricated world overriding lived action.” Another example is as follows: (i) Raw excerpt: “We fill in the platform because we must, but no one believes these numbers capture anything real.”; (ii) Inductive code: “symbolic compliance / disbelief”; (iii) Linked Arendtian concept: “thoughtlessness (mechanical conformity)”; (iv) Theoretical note: “Resonates with literature on performativity and metric distortion”.

Notably, sensitising concepts (see Table 2, first column) were not imposed as fixed coding structures from the outset but were used as analytical lenses to interpret patterns emerging from the data. Their use is grounded in the theoretical framework, which provides a conceptual vocabulary for understanding different modes of activity and experience. In this sense, these categories do not replace empirical themes but offer a way of organising and interpreting them in relation to the study’s theoretical concerns. All supporting evidence reported in Table 2 is drawn primarily from interview data, complemented by insights from autoethnographic reflections and institutional materials if relevant.

Table 2. Coding

Level 1 – Sensitising concept derived from the framework	Level 2 – Empirical codes	Concepts emerging from the interviews
Labour (repetition, necessity, consumption)	Labour-like repetitive tasks	Cyclical reporting, annual re-uploading of activities, continuous clarification of “what counts”

	Bureaucratic overload	Administrative burden associated with platforms, templates, and documentation
	Invisible or unrecognised labour	Engagement work not acknowledged in careers, workload models, or incentives
	Unequal distribution of labour	Disproportionate burden on women, junior academics, and delegates
	Fatigue and depletion	Emotional and cognitive exhaustion linked to compliance demands
Work (institutional artefacts, world-building)	Institutional artefacts as work	VQR, AVA, IRIS, SUA-TM, templates and platforms as stabilising devices
	Definitional ambiguity	Persistent uncertainty about boundaries and categories of Third Mission
	Category mismatch frustration	Difficulty fitting meaningful activities into predefined evaluative categories
	Visibility-oriented structuring	Privileging activities that are countable, scalable, and communicable
	Stakeholder selectivity	Preference for “safe”, visible, or well-resourced partners
	Marginalisation of relational activities	Exclusion of slow, trust-based or culturally embedded engagement
Action (plurality, relationality, initiative)	Relational engagement as value	Emphasis on trust, co-creation, long-term relationships
	Desire for meaningful engagement	Aspirations to go beyond compliance toward socially relevant action
	Small-scale but high-impact work	Engagement with schools, NGOs, grassroots groups
	Action constrained by evaluation	Limitation of initiatives due to metric requirements
	Shrinking spaces for dialogue	Lack of genuine arenas for co-decision and plurality
Judgment / Thinking (reflection, ethical evaluation)	Critical evaluation of metrics	Explicit questioning of indicator-based definitions of value
	Ethical responsibility to communities	Commitment to societal needs over institutional visibility
	Identity reflexivity	Reflection on expanded or fragmented academic roles
	Recognition injustice	Perceived mismatch between effort, value, and rewards
	Normative critique of governance	Explicit judgments about unfairness, misalignment, or distortion
Thoughtlessness (procedural conformity)	Symbolic compliance	Filling platforms “because we must”, without belief in their meaning
	Mechanical conformity	Following procedures despite recognising their inadequacy
	Steered autonomy	Formal freedom combined with substantive constraint
	Performative participation	Consultation without influence on decisions

	Translation and adaptation strategies	Reshaping projects to improve traceability rather than substance
	Dual accountability tension	Conflict between bureaucratic accountability and relational responsibility

This iterative coding process was accompanied by regular reflexive discussions among the research team members to calibrate interpretations, challenge assumptions, and ensure analytic consistency. Such an approach allowed the authors also to manage any auto-reflexive bias related to the autoethnographic component. Notably, level 1 categories were not used to exhaustively classify data, but rather to guide interpretation and theorisation. Level 2 codes often appeared across multiple level 1 categories, reflecting the entanglement of labour, work, action, judgment, and thoughtlessness in practice. The analytical contribution lies not in the catalogue itself, but in tracing how evaluative infrastructures shift academic engagement from action toward labour and procedural conformity. The role of the researchers as insiders was explicitly acknowledged throughout data interpretation, consistent with Haynes’ (2011) emphasis on reflexivity as a methodological asset rather than a liability. Rather than seeking a notion of objectivity premised on distance, the analysis draws from the dialogic encounter between personal experience and participants’ narratives to illuminate tensions that may remain invisible in more detached research designs.

Altogether, this methodological approach provides a robust empirical foundation for understanding how academics experience the expanding evaluative infrastructures of the Italian Third Mission, and how these infrastructures shape the balance between repetition, structure, and meaningful engagement. By mobilising Arendt’s conceptual distinctions in conjunction with rich empirical material, the methodology enables a nuanced exploration of how evaluation systems intersect with academic agency, identity, and the possibility for action within the contemporary university.

5. Setting the Context: The Italian Third Mission

This study examines the accountability implications and engagement challenges surrounding academics’ roles in the Third Mission within Italian universities. It focuses on how *the metrification of TMA experienced by academics*, with particular attention to how performance evaluation systems shape academic agency, identity, and conceptions of societal responsibility. Understanding these dynamics requires situating the inquiry within the broader policy environment that structures the Italian higher education system, an environment characterised by dense regulation, a strong evaluative state, and longstanding tensions between managerial control and academic autonomy.

Across Europe, higher education has been substantially reconfigured by NPM, which promotes efficiency, accountability, and performance as guiding principles of governance (Deem & Brehony, 2005) and by increasing requirements for universities to legitimise themselves not only for knowledge generation through research and dissemination through teaching, but also for the “value” they generate for society. The expansion of the so-called Third Mission reflects this shift, encompassing public engagement, knowledge transfer, social innovation, lifelong learning, and cultural and scientific dissemination, with an impact on governance and commercialization of university’s outputs (Parker, 2024). While often presented as an opportunity to enhance universities’ social relevance, the Third Mission has increasingly been inserted into managerial architectures that privilege standardisation and comparability.

Italy represents a particularly dense and revealing case of these developments because it makes especially visible the broader transformation of societal engagement into an object of governance through performance evaluation. The Italian higher education system combines a strong national

regulatory framework with an increasingly elaborate architecture of quality assurance and research assessment, allowing the dynamics associated with the metrification of the Third Mission to emerge with particular clarity. Over the past two decades, the Third Mission has been progressively institutionalised through a series of legal, administrative, and evaluative reforms, producing a complex landscape in which societal engagement is both institutionally mandated and bureaucratically mediated. A pivotal moment was the Gelmini Reform (Law 240/2010), which established the AVA system (Self-Evaluation, Periodic Evaluation, and Accreditation) and explicitly recognised universities' interactions with their socio-economic context. This reform laid the foundation for the structured governance of Third Mission activities. The subsequent creation of the Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR) in 2011 consolidated the evaluative orientation of the sector by centralising authority over assessment practices, including those tied to the Third Mission. Legislative decree 19/2012 and ministerial decree 47/2013 further codified these activities through tools such as the SUA-TM (Scheda Unica Annuale della Terza Missione – Single Annual Report on Third Mission, a standardized reporting framework through which Italian universities documented and reported their Third Mission and societal impact activities for monitoring and evaluation purposes), transforming what were once heterogeneous, locally defined forms of engagement into objects of audit and oversight.

Successive reforms deepened and broadened this evaluative architecture. AVA 2.0 (2016) and ANVUR's 2015 Manual expanded the categories of recognised Third Mission activities, e.g. public engagement, patents, cultural heritage, spin-offs, and linked them explicitly to national research assessment exercises (VQR). By 2018, universities were required to integrate Third Mission objectives into institutional and departmental strategic plans, marking a shift from voluntary engagement to formalised responsibility embedded in resource allocation mechanisms. More recently, AVA 3.0 (Ministerial Decree 1154/2021) created a dedicated evaluation area for research quality and Third Mission/social impact, signalling that societal engagement had become not only symbolically important but materially consequential for institutional positioning.

This regulatory trajectory has redrawn the boundaries within which Italian universities operate. While policies officially frame the Third Mission as a vehicle for enhancing societal value, they simultaneously bind it to bureaucratic procedures, standardised metrics, and funding incentives. These mechanisms do not merely document existing practices; they define and delimit what counts as legitimate engagement. The VQR, for instance, tends to privilege activities producing measurable and visible outputs, creating asymmetries that benefit disciplines with quantifiable or commercialisable forms of impact, often at the expense of humanities or social sciences. Likewise, the procedural requirements associated with AVA 3.0 introduce uniform templates that sit uneasily with the diversity of disciplinary traditions and institutional missions, resulting in administrative burdens that unevenly affect departments (Spanò et al., 2022).

The boundaries of the Italian evaluative system are therefore marked by a persistent tension between intrinsic and extrinsic rationales for engagement. On the one hand, the Third Mission is promoted as an ethical, relational, and context-dependent responsibility aligned with universities' public purpose. On the other hand, its implementation is filtered through audit logics that privilege compliance, traceability, and comparability (Benneworth & Jongbloed, 2010). This duality constrains the space for locally meaningful forms of engagement and reinforces a mode of governance in which indicator-based performance becomes the primary language through which societal contribution is expressed.

Within this environment, academics are not mere passive recipients of policy but key actors navigating the potential contradictions introduced by metrification. They encounter expanding expectations for societal engagement while simultaneously working within structures that contribute to shape the meaning of what engagement entails. Yet despite the centrality of academics to the

enactment of the Third Mission, their perspectives on how evaluative systems affect their agency, identities, and everyday practices remain relatively unexplored. This gap provides the empirical rationale for the present study. Understanding how academics experience Third Mission evaluation also requires brief consideration of academic contracts and incentives in Italy. While teaching hours are formally defined in employment contracts, activities related to research, administration, and societal engagement are not subject to clear workload allocation. Research performance is strongly incentivised through national assessment exercises such as the VQR, which directly affect careers and institutional resources. By contrast, Third Mission activities are formally evaluated (under AVA3 and in VQR exercises) but remain weakly integrated into promotion criteria and workload models. As a result, engagement often becomes a necessary but insufficient condition for career advancement, shaped more by visibility and informal power dynamics than by explicit contractual recognition. This institutional configuration is crucial for interpreting why Third Mission work is frequently experienced as unrecognised labour.

6. Findings: An Arendtian Interpretation of Academic Experiences of the Third Mission

The empirical material highlights how the institutionalisation of the Third Mission reconfigures academic work in ways that strongly resonate with Arendt's distinctions between labour, work, and action, as well as her reflections on thinking, judgment and thoughtlessness. Across interviews and autoethnographic notes, participants describe environments characterised by repetitive and often invisible tasks, by rigid evaluative artefacts that shape what counts as value, by shrinking spaces for relational engagement, and by ambivalent forms of autonomy in which critical judgment is both exercised and eroded.

The findings are organised using Arendt's categories. For each category, we show how participants' experiences reflect the dynamics described in the theoretical framework and how these dynamics intersect in shaping the lived reality of the Third Mission. Rather than treating the categories as fixed labels for sorting data, the analysis uses them as flexible, relational lenses. This makes it possible to see how academic experiences shift across - and often combine - different modes of activity, such as repetition and initiative, compliance and critique, or relational engagement and procedural constraint. By following these tensions and movements, the analysis moves beyond listing practices to examine how evaluative systems actively reshape the conditions for judgment, responsibility, and meaningful engagement. In this way, the categories are used not to stabilise empirical realities, but to question them, drawing attention to what is constrained, displaced, or made fragile within metric-driven regimes.

6.1 Labour: Repetition, Overload and Unequally Distributed Effort

All of the participants contended how Third Mission activities that were previously part of ordinary academic life - such as public lectures, collaborations with schools, cultural initiatives, or long-term relationships with local organisations - have become increasingly shaped by detailed reporting requirements. While these activities themselves are often experienced as meaningful, the need to document and classify them within institutional systems introduces a new layer of work that is repetitive, time-consuming, and often perceived as disconnected from the substance of engagement. As UD1 notes, many colleagues were already deeply engaged with society, yet "*this effort was rarely visible or institutionally rewarded*". Formal recognition does not eliminate this invisibility; instead, it introduces new cycles of documentation that must be repeated annually under current norms and internal consolidated approaches, often according to shifting criteria. Rather than producing a durable institutional achievement, recognition itself becomes contingent on its continual reproduction, suggesting that visibility is experienced as an ongoing requirement rather than a lasting accomplishment.

A central issue emerging from the data concerns the persistent uncertainty around “*what counts*” as Third Mission. DD1 describes being repeatedly approached by colleagues seeking clarification about whether their activities fit institutional definitions. These requests require ongoing interpretation, as evaluative categories are not stable but evolve over time, depending on changes in platforms, templates, and national guidelines. As a result, “*the same activities must be continuously explained, justified, and reframed*” (autoethnographic notes). Participants, such as UD2 emphasise that “*this work does not accumulate into lasting recognition but must be constantly reproduced*”. On this regard, UD3 claimed that this “*tends to generate a sense of circularity in which effort is repeatedly expanded without producing durable outcomes*”. Rather than reducing ambiguity, the evaluative system continually regenerates it, requiring academics to repeatedly reinterpret and rearticulate the meaning of their activities in response to shifting institutional expectations. Consequently, Third Mission work becomes characterised by an ongoing cycle of interpretive labour in which recognition is provisional rather than cumulative, and value must be continuously reconstructed rather than simply demonstrated. Participants therefore describe engagement not as a completed accomplishment but as something that must be continually renewed through recurrent administrative effort, echoing the cyclical quality that characterises labour.

Crucially, most interviewees distinguish between the meaning they attribute to engagement and the effort required to make it visible within evaluative systems. P1 observes that Third Mission activities are “*essential but invisible*”, while R3 describes them “*as marginal to the ‘real’ academic core of research and teaching*”. These accounts suggest that the effort associated with documenting engagement is not experienced as intrinsically valuable, but as a requirement imposed by institutional expectations. From this perspective, as corroborated by the authors’ autoethnographic experience, and by claims by D1 “*the burden does not originate in engagement itself, but in the need to translate it into measurable and comparable forms*”.

This burden is also unevenly distributed across the academic workforce. Interviewees and autoethnographic notes converge on the observation that a relatively small group of actors - often women, junior academics, and those holding formal or informal coordination roles - carry the bulk of reporting, organisational, and interpretive work. As DD1 remarks, she is “*asked to lead, motivate, and report, but without formal acknowledgement of the time and energy required*”. Others note that while many colleagues participate in moments of public visibility, “*only a few engage in the ongoing ‘dirty work’ required to sustain the evaluative apparatus*” (D2). These patterns point to structural asymmetries in how the demands of evaluation are allocated and experienced. Labour is therefore not only repetitive but also unevenly allocated, with its most routine and maintenance-oriented dimensions becoming concentrated among specific academic groups.

From a processual perspective, the majority of participants describe a shift in how engagement unfolds over time. When engagement develops organically, meaningful interaction typically precedes its formalisation. Under metric-driven regimes, however, “*documentation increasingly comes first*” (D3). Academics report that they are required to anticipate how activities will be recorded and assessed, often “*reshaping or prioritising them based on their traceability rather than their societal relevance*” (P2). This reorientation contributes to feelings of fatigue, frustration, and depletion, as time and cognitive resources are absorbed by activities that do not generate a corresponding sense of purpose or recognition.

These patterns are interpreted through Arendt’s notion of labour, understood as cyclical, necessary activity whose outputs are immediately consumed and must be continuously reproduced without leaving a durable trace. In this case, labour does not reside in engagement itself, but in the ongoing effort required to render it visible within evaluative systems. The repetitive documentation, shifting criteria, and lack of cumulative recognition reflect key features of labour: dependence on necessity,

absence of lasting outcomes, and continuous reproduction. More broadly, these dynamics illustrate Arendt's concern with conditions in which actors become structurally bound to necessity, limiting their capacity to engage in more meaningful forms of activity.

6.2 Work: Evaluative Artefacts and the Construction of What "Counts"

Across almost all interviews, participants describe how Third Mission engagement is increasingly shaped by a range of institutional artefacts, including templates, databases, reporting platforms, classificatory schemes, and strategic documents. These tools are not experienced as neutral instruments for documenting activities, but as structures that actively define how engagement should be represented and organised. As UD3 notes, while the growing attention to societal engagement is welcomed in principle, *"the boundaries of the Third Mission remain difficult to define"*, and even when they appear clearer, *"questions of evaluation persist"*. Similarly, an early-career participant (R2) reflects that *"you spend more time understanding how to frame the activity than actually doing it"*, pointing to the practical implications of these ambiguities in everyday work. This suggests that attempts to formalise engagement through institutional tools do not eliminate ambiguity but instead reconfigure how it is experienced. Participants portray these artefacts as giving stable form to an otherwise fluid domain, transforming engagement into something that can be organised, classified, and preserved within institutional structures.

A recurring theme in the data is that these artefacts shape not only how engagement is reported, but *what engagement is recognised as legitimate* (autoethnographic critical note). Participants describe ongoing uncertainty around *"what counts"* as Third Mission, requiring continuous interpretation and negotiation. DD2, for example, refers to persistent ambiguity about *"what exactly counts as Third Mission,"* noting that this *"muddies value criteria"*. Rather than stabilising meaning, evaluative systems appear to generate *"new forms of uncertainty"*, requiring academics to constantly align their activities with shifting definitions and expectations. This sense of instability is echoed across roles, with one participant noting that *"just when you think you understood the criteria, they change again"* (DD3), highlighting how uncertainty becomes embedded in routine practice rather than resolved over time. The search for stability therefore depends less on shared interpretations than on increasingly durable classificatory devices.

Several interviewees explicitly describe how evaluative artefacts operate as gatekeeping devices. R3 explains that *"certified systems rule the whole problem: if something is not already categorised, it cannot be reported... even if it has a potentially relevant impact for society"*. In this sense, institutional infrastructures do not merely record engagement; they actively determine what becomes visible and legitimate within the system. Activities that cannot be translated into predefined categories remain effectively excluded, regardless of their societal value. This reinforces the perception that recognition is less about substance and more about alignment with predefined formats. In participants' accounts, institutional artefacts acquire an organising capacity that exceeds documentation, becoming the very medium through which engagement is rendered real and institutionally actionable.

This selective visibility extends to relationships with external partners. Participants report pressures to prioritise collaborations that are easily communicable, standardised, and aligned with institutional expectations. R1 notes that *"grassroots feminist groups are often perceived as too political"*, while UD2 and DD2 describe a preference for initiatives that *"look good on paper."* Similarly, D1 and D2 observe that smaller NGOs or community organisations are frequently excluded, not because their work lacks value, but because it does not fit existing evaluative templates. Few interviews also reported stories of colleagues actively avoiding certain collaborations because they were perceived as *"difficult to translate into the system"* (DD3), suggesting that evaluative constraints not only filter visibility ex post, but shape engagement choices ex ante. Rather than simply recording action, these

infrastructures progressively configure the conditions under which action can be imagined and undertaken.

Importantly, participants do not attribute these dynamics to individual decisions alone, but to the design and durability of evaluative systems. This is also reflected in autoethnographic notes, which document how even minor decisions about engagement activities are increasingly filtered through questions such as “*how will this be reported?*” or “*under which category will this fall?*”. Once templates, indicators, and platforms are established, they structure how engagement can be conceived and enacted. Academics may be aware of the limitations of these tools, yet still feel compelled to operate within them, as deviation risks invisibility or non-recognition. From a temporal perspective, this leads to a shift in how engagement is organised: rather than emerging from prior relationships or initiatives, activities are increasingly designed to satisfy reporting requirements and formal compliance, with their substantive societal value becoming secondary to their evaluative traceability. Anticipation of evaluation becomes a starting point, shaping what is pursued and how it is structured. Autoethnographic reflections further highlight how this anticipatory orientation becomes habitual, shaping planning practices even before engagement takes place.

These patterns can be interpreted through Arendt’s notion of work, which refers to the fabrication of relatively durable artefacts that give structure and stability to human activity. The empirical material shows how evaluative infrastructures perform precisely this function, defining the conditions under which engagement becomes visible and meaningful within the institution. At the same time, this fabricated world imposes boundaries: it enables coordination and comparability, but also constrains what can be recognised, valued, and pursued. In this sense, evaluative work not only supports institutional organisation, but also shapes the horizon of possible action, reinforcing the tendency for engagement to be aligned with existing categories rather than emerging from open-ended, relational processes.

6.3 Action: Relational Value and the Narrowing of the Space for Plural Engagement

A large majority of the participants describes societal engagement as a relational and meaningful dimension of academic work, grounded in interaction, trust, and long-term commitment to communities. Across interviews, Third Mission activities are not framed primarily in terms of measurable outputs, but as processes that unfold through sustained relationships and responsiveness to societal needs. As D3 argues, value is “*not only about economic return or measurable outputs; it’s about creating conditions where people can thrive, communities can trust us, and the knowledge we produce serves society in meaningful ways*”. Similarly, DD1 emphasises that “*true value lies in long-term relationships and slow processes of trust-building*,” while R2 associates meaningful engagement with “*ethical commitments and relevance to real societal needs*”. These accounts indicate that participants share a clear and articulated understanding of engagement as a relational and plural practice. Accordingly, the value of engagement emerges through interaction with others rather than through predefined outputs, emphasising its inherently relational character.

This orientation toward meaningful engagement is echoed across different roles and levels of seniority. For instance, one participant notes that “*the most important work we do with society is often the least visible*” (R3), highlighting the tension between perceived value and institutional recognition. In a similar vein, autoethnographic notes document how some of the most impactful interactions with external partners - such as informal exchanges, mentoring relationships, or community-based initiatives - rarely enter formal reporting systems, despite being central to how academics understand their societal role. These practices derive their significance from the relationships they sustain rather than from their capacity to generate measurable outputs.

At the same time, participants provide details to describe institutional conditions that constrain their capacity to engage in these ways. These constraints do not stem from a lack of motivation or awareness, but as some of them further clarified, from the growing dominance of evaluative structures. As reporting requirements intensify and evaluative artefacts define what counts as legitimate engagement, academics experience increasing pressure to prioritise activities that are visible, standardisable, and low risk. R1 notes that *“small-scale work with schools or advocacy groups, often the most impactful, gets overlooked because it does not generate impressive numbers”*. DD1 similarly observes that *“impactful but unglamorous work is marginalised, while visibility-driven initiatives are privileged”*. Several participants describe actively reshaping or limiting their engagement choices to fit evaluative expectations, even when these diverge from their own understanding of societal value. As a consequence, initiative increasingly becomes oriented towards institutional recognisability rather than the unpredictable development of relationships with external communities.

This dynamic does not eliminate engagement but selectively reshapes it. Activities that align with institutional templates remain possible, while forms of engagement that depend on slow interaction, political sensitivity, or locally embedded trust become increasingly difficult to sustain. Participants describe a narrowing of acceptable forms of engagement, where relational and open-ended initiatives are either adapted to fit evaluative formats or pushed to the margins. As one participant notes, *“you can still do meaningful work, but you have to fit it into the system, otherwise it doesn’t exist”* (P2). This suggests that visibility within evaluative frameworks becomes a precondition for recognition, rather than a consequence of meaningful engagement. Taken together, these accounts suggest that participants experienced engagement as progressively losing the openness that they regard as essential to its societal value.

Participants also highlight how the space for engagement is constrained by limited opportunities for genuine participation in decision-making. DD1 remarks that *“very few forums qualify as true dialogue”*, adding that academics are often consulted only *“to validate predefined decisions”*. R1 echoes this experience, stating that *“when you are not in a leadership position, the system treats you as a data provider.”* P1 summarises this condition as participation that is *“more formality than genuine co-decision”*. These accounts indicate that the capacity to shape the meaning and direction of the Third Mission is unevenly distributed, further limiting opportunities for collective and dialogic forms of engagement.

Importantly, the findings do not point to a disappearance of meaningful engagement, but to its increasing fragility. Participants describe continuing to invest in relationships with communities, often *“under the radar,”* sustaining forms of engagement that are not fully captured by evaluative systems. However, this often comes at the cost of additional labour, strategic compromise, or reduced visibility. Autoethnographic reflections similarly document how engagement practices are frequently adjusted to balance institutional expectations with personal commitments, resulting in a continuous negotiation between what is meaningful and what is measurable.

These patterns can be interpreted through Arendt’s notion of action, understood as the sphere of plurality, initiative, and relational engagement through which individuals contribute to a shared public world. Action is inherently situated, unpredictable, and dependent on interaction with others. The empirical material shows that while academics retain a strong orientation toward such forms of engagement, the institutional conditions that sustain action are increasingly constrained. Evaluative regimes oriented toward control, comparability, and visibility sit uneasily with the open-ended and relational character of action, narrowing the space in which it can occur. In this sense, action does not disappear, but becomes fragmented and precarious, sustained more by individual commitment than

by institutional support, and continuously negotiated within the limits imposed by metric-driven governance.

6.4 Judgment: Critical Evaluations of Value, Equity and Identity

All the interviewees consistently demonstrate a strong capacity to reflect critically on how societal engagement is defined, valued, and organised within their institutions. Rather than passively accepting evaluative frameworks, interviewees articulate nuanced and often critical assessments of what counts as meaningful engagement, how value is attributed, and whether institutional practices align with declared goals. This suggests that academics do not lack awareness or normative orientation but actively question the coherence and adequacy of existing arrangements. Their accounts are characterised not simply by dissatisfaction but by sustained reflection on alternative ways of recognising and organising societal engagement.

A majority of the participants highlight tensions between their own understanding of societal value and the criteria embedded in evaluation systems. R2 explicitly contrasts their view of value, grounded in *“long-term relationships, ethical commitments, and relevance to real needs of society”*, with what they describe as the narrower criteria adopted by evaluation agencies. This is not merely an observation of difference, but a critical evaluation of inadequacy. Similar reflections emerge across roles, with one participant noting that *“the system captures what is visible, not necessarily what is valuable”*, (D1) pointing to a perceived gap between measurement and meaning. Rather than simply describing institutional evaluation criteria, several participants critically questioned their underlying assumptions, fairness, and ability to capture meaningful societal engagement.

Critical reflections are particularly evident in discussions of equity and recognition. DD2 highlights how *“certain forms of engagement, particularly those related to gender, care, or social inequality, continue to be perceived as ‘soft’ and insufficiently legitimised within institutional hierarchies of value”*. These assessments reveal an awareness of the normative assumptions embedded in evaluative frameworks and a willingness to question them. Rather than accepting existing hierarchies as given, participants identify them as partial, selective, and potentially unjust. These reflections reveal that participants do not treat evaluative categories as neutral but question the assumptions on which they are built.

Judgment also emerges in participants’ evaluations of organisational coherence and incentives. Several interviewees point out that while Third Mission engagement is rhetorically promoted, it remains weakly integrated into promotion criteria, workload allocation, and funding structures. UD1 notes that *“visibility may be celebrated, but time credits and funding lines are missing”*, while DD2 observes that existing incentives are *“symbolic rather than structural”*. These reflections converge on a shared critique: institutions encourage expanded societal engagement without providing the material conditions necessary to sustain it. From this perspective, participants are not simply describing constraints but actively evaluating them as inconsistent with institutional rhetoric.

Reflections on academic identity further illustrate this critical capacity. P2 explains that she no longer sees herself solely as a researcher or professor, but also as *“a facilitator, mediator, and advocate”*. R1 similarly remarks that academics are increasingly expected to be *“communicators and network builders”* yet feel *“stretched too thin”*. These accounts show that participants are actively interpreting and assessing the transformation of their professional roles, recognising both its potential value and its practical limits. Autoethnographic notes reinforce this point, documenting moments in which engagement activities are experienced as meaningful yet difficult to sustain within existing institutional structures, requiring ongoing negotiation between personal commitment and formal

expectations. Professional identities are therefore not simply changing but are actively interpreted and reassessed by academics themselves.

These patterns can be interpreted through Arendt's notion of judgment, understood as the capacity to think reflectively about what is right, meaningful, or appropriate beyond the application of predefined rules. Judgment enables individuals to step back from established standards and evaluate situations in their particularity. The empirical material shows that such judgment is not absent; rather, it is actively exercised across multiple domains, including value, equity, identity, and institutional coherence. However, its effects remain largely confined to the individual level.

In this sense, the issue is not the absence of judgment, but its disconnection from collective and institutional arenas. While academics continue to reflect critically on their practices and contexts, the structures within which decisions are made provide limited space for such reflections to shape outcomes. Judgment persists as an internal and often shared awareness but struggles to translate into coordinated action or institutional transformation. This disconnection is central to understanding the dynamics of the Third Mission: it is not that academics fail to think critically, but that the conditions under which judgment can inform practice are progressively constrained. As a result, reflective capacity endures, yet its practical and political significance is diminished, creating the conditions under which compliance can become routinised and thoughtlessness gradually normalised.

6.5 Thoughtlessness: Procedural Compliance, Steered Autonomy and Dual Accountability

A vast majority of the participants describes a mode of engagement increasingly characterised by procedural compliance. While academics retain formal autonomy over their activities, they experience growing constraints linked to what can be monitored, reported, and evaluated. DD1 notes that dialogue around the Third Mission is "*mainly top-down*," while autonomy exists "*on the surface*," constrained by reporting requirements and institutional expectations. Similarly, R1 remarks that "*discretion is formally preserved but substantively narrowed*", as engagement becomes shaped by the need to align with evaluative frameworks. Participants distinguish between having formal discretion and experiencing genuine autonomy in practice, suggesting that compliance increasingly structures everyday decision-making.

A recurring pattern in the data is the emergence of what participants describe as symbolic compliance. Activities are reported because "*we must*," rather than because reporting is perceived as meaningful or informative. As one interviewee observes, "*we fill in the platform, but no one believes these numbers capture anything real*" (D3). This does not reflect disengagement from societal impact, but rather a disjunction between internal convictions and external requirements. Academics continue to value meaningful engagement yet participate in evaluative routines that they perceive as only loosely connected to that purpose. Procedural completion gradually becomes an end in itself, independently of whether participants believe that reporting captures the substance of engagement.

This disjunction is reinforced by the experience of dual accountability. Most participants describe being accountable externally to societal partners through relationships of trust, reciprocity, and ethical commitment, while internally they are accountable to deadlines, templates, and performance indicators. UD1 explains that a collaboration may be successful and meaningful for a community, yet "*count for little if it does not produce a metric*". DD1 articulates this tension explicitly: "*I am accountable to partners through trust, but internally to forms and deadlines. These logics clash*." UD3 warns also that "*in situations of conflict, institutional accountability tends to prevail, not because it is perceived as more legitimate, but because it carries formal consequences*". Taken together, these accounts suggest that academics navigate two competing normative orders: one grounded in relational responsibility towards external communities and another centred on procedural

accountability to institutional evaluation systems. As only the latter is formally recognised and rewarded, several participants described experiencing persistent tensions in deciding where to invest their limited time and effort, often feeling compelled to prioritise activities that satisfied institutional reporting requirements over forms of engagement they regarded as more socially meaningful.

Over time, these pressures encourage adaptive strategies that reshape engagement practices. Several participants reported adapting their projects to make them more easily traceable within institutional evaluation systems, for example by merging initiatives into existing reporting categories or prioritising activities that were easier to document. P1 notes that meaningful engagement often proceeds “*with no recognition*”, while DD3 explains that projects are sometimes reconfigured “*to make them visible rather than better*”. These strategies are not experienced as imposed externally, but as necessary adjustments to operate within the system. Autoethnographic notes further illustrate how this adaptation becomes routine, with decisions about engagement increasingly guided by anticipatory considerations such as visibility, categorisation, and evaluability. These adjustments are rarely experienced as exceptional compromises but as ordinary ways of working within the evaluative environment.

As these patterns become embedded in everyday practice, a gradual shift occurs in how academics relate to evaluative processes. While participants continue to express critical awareness of the limitations of existing systems, these reflections are increasingly decoupled from action. The question of whether indicators capture what truly matters does not disappear but becomes secondary to the practical need to comply with established procedures. In this sense, reflective capacity persists, yet its influence on practice is diminished.

These dynamics can be understood through Arendt’s notion of thoughtlessness, which does not imply a lack of intelligence or moral concern, but the suspension of reflective judgment in favour of routine, rule-following behaviour. Thoughtlessness emerges in contexts where institutional structures reward compliance, discourage questioning, and separate thinking from action. The empirical material shows that such conditions are progressively produced through the institutionalisation of the Third Mission: academics remain capable of judgment but are increasingly oriented toward procedural alignment as the primary condition for recognition. Importantly, thoughtlessness in this context is neither total nor uncontested. It coexists with judgment, ethical concern, and genuine commitment to societal engagement, as shown in the previous subsection. However, as evaluative infrastructures expand and stabilise, they increasingly channel academic behaviour toward routinised compliance, narrowing the space in which reflective judgment can inform action. In this sense, thoughtlessness emerges not as an individual failure, but as a structural outcome of metric-driven governance.

Taken together, this completes the process traced across the findings. Engagement initially grounded in relational action becomes progressively shaped by evaluative artefacts (work), then absorbed into repetitive reporting cycles (labour), and ultimately reoriented toward procedural compliance (thoughtlessness). This transformation does not occur through coercion, but through the ordinary functioning of evaluative systems, which gradually redefine what counts as legitimate academic engagement.

7. Discussion

This study examined how the metrification of Third Mission activities is lived, interpreted, negotiated, and sometimes resisted by internal stakeholders within Italian universities. Building on an Arendtian framework, the analysis goes beyond well-established critiques of performance management in higher education concerning the increasing reliance on quantitative and judgement-oriented evaluation systems (ter Bogt & Scapens, 2012), the performative consequences of metric-

based evaluation (Kallio et al., 2017a; Gendron & Rodrigue, 2021), the managerialisation and narrowing of academic value (Argento & van Helden, 2023; Carnegie, 2022, 2023), and the organisational burdens associated with expanding performance regimes (Aboubichr & Conway, 2023). Rather than treating evaluation as a technical overlay on academic work, the findings show how evaluative infrastructures reshape the very conditions under which societal engagement acquires meaning, legitimacy, and direction.

Existing literature has shown that metric regimes privilege what is measurable over what is meaningful, often narrowing academic priorities and encouraging symbolic compliance (Michelon, 2021; Gebreiter, 2022; Parker, 2024). Our findings confirm these patterns in the context of the Third Mission, but extend them by revealing the *process* through which this narrowing occurs. The analysis showed that central process that emerged across interviews and autoethnographic reflections. Most participants described a reversal in the sequence through which societal engagement is enacted and valued. When Third Mission activities develop organically, engagement typically begins with *action* (relational, meaningful interaction with societal actors), which may later generate durable outputs (*work*) and, eventually, routine tasks of documentation (*labour*). Under metric-driven evaluation regimes, this sequence is inverted. Engagement increasingly begins with evaluative artefacts, templates, platforms, indicators, that define in advance what is expected, visible, and legitimate. Labour in the form of repetitive reporting follows, while opportunities for action are constrained to those forms already anticipated by the evaluative system. This inversion provides a processual explanation for why Third Mission evaluation so often generates fatigue, frustration, and a perceived loss of meaning, echoing broader critiques of audit cultures in higher education (Deem & Brehony, 2005; Manes Rossi et al., 2022; Spanò et al., 2024). Such evidence helps explain why evaluative systems risk fostering procedural compliance and thoughtlessness, rather than meaningful societal engagement.

Arendt's distinction between labour, work, and action helps clarify the implications of this transformation. Much of the existing literature documents the growth of administrative workload associated with accountability systems (Kallio et al., 2017b; Spanò et al., 2024). Our findings show that this labour is not simply an unintended side effect, but a structural outcome of evaluative regimes that require engagement to be continuously translated into measurable and comparable forms. At the same time, evaluative infrastructures operate as forms of *work* in the Arendtian sense: they fabricate a durable institutional world that stabilises definitions of societal value. Templates, databases, and indicators do not merely record engagement; they shape what can appear as engagement in the first place, privileging visible, scalable, and low-risk activities while marginalising slow, relational, or politically sensitive forms of interaction. This world-building function resonates with prior studies on the institutionalisation of impact and knowledge transfer (Benneworth & Jongbloed, 2010), but the Arendtian lens reveals its deeper implications for academic agency and meaning.

Importantly, the findings do not suggest that academics lack a sense of societal responsibility or relational commitment. On the contrary, participants articulate understandings of value grounded in trust, ethical responsibility, and long-term engagement, consistent with dialogic and stakeholder-oriented perspectives (Morsing & Schultz, 2006; Steccolini, 2019). What changes is not the presence of action itself, but the institutional conditions within universities that enable and sustain it. Action becomes fragile, selective, and often dependent on individual commitment rather than organisational support. Engagement persists, but increasingly, is situated “under the radar,” as academics seek to protect relational practices from evaluative distortion.

A key contribution of this study lies in showing that the expansion of evaluative regimes does not eliminate judgment but displaces it. Academics repeatedly exercise critical judgment regarding the adequacy of indicators, the inequitable distribution of labour, and the misalignment between

institutional rhetoric and material incentives. These judgments resonate with broader debates on plural value creation and public accountability (Bracci et al., 2021; Sicilia et al., 2025). Yet they remain weakly connected to organisational decision-making. Consultation processes are often experienced as formalistic, reinforcing patterns of performative participation observed in other public sector contexts. As a result, judgment becomes individualised and politically muted, rather than collectively transformative.

This disconnection is intensified by conditions of dual accountability. Academics are accountable externally to societal partners through relationships of trust and reciprocity, and internally to indicators, platforms, and deadlines. Similar tensions have been observed in public sector organisations, where bureaucratic accountability can undermine relational forms of value creation (Steccolini, 2018; Tanima et al., 2024). In the university context, however, these tensions are exacerbated by the weak integration of Third Mission activities into formal workload models and career incentives (Manes Rossi et al., 2022). When conflicts arise, metric-based accountability tends to prevail, not because it is perceived as more legitimate, since quantified indicators often acquire authority not because they capture value more adequately, but because they make complex activities visible, commensurable, and governable (Kallio et al., 2017a; Spanò et al., 2024). In the context of Third Mission evaluation, this means that relational responsibility to communities may be subordinated to forms of accountability that are easier to document, verify, and defend within institutional procedures.

Read through Arendt's reflections (Arendt, 1963; 1978), these dynamics point to a systemic risk. Thoughtlessness here does not imply moral indifference or lack of critical capacity. Rather, it emerges when evaluative routines normalise the separation between thinking and action. Academics may continue to reflect critically on what constitutes meaningful societal engagement yet find themselves enacting procedures they privately question. Over time, compliance becomes habitual and questions about whether indicators capture what matters recede, making thoughtlessness an institutional outcome rather than an individual failure.

By bringing these elements together, the study contributes to the literature on performance management in higher education in three main ways. First, it offers a processual account of how metrification reshapes societal engagement, identifying the inversion of the action-work-labour sequence as a central mechanism through which meaning is eroded. Second, it shows that resistance and compliance coexist: academics may critically judge evaluative regimes while simultaneously reproducing them through routine practices. Third, it reframes thoughtlessness as a structural condition produced by evaluative infrastructures, redirecting attention from individual behaviour to the institutional conditions that enable or constrain thinking, plurality, and action. In doing so, the paper complements existing research on metric regimes and public value by shifting the focus from indicators and incentives to the deeper organisational dynamics that shape how universities enact their societal role.

8. Conclusions

The objective of this study was to understand how academics experience the metrification of Third Mission activities and how these evaluative processes reshape the conditions for labour, work, and action within the contemporary university. Drawing on an Arendtian perspective and the experiences of academics operating within the Italian higher education system, the analysis shows that, far from

representing a straightforward extension of academic responsibility, the Third Mission has become a complex and often contradictory terrain where aspirations for societal engagement collide with the constraints of evaluative regimes. Through Arendt's vocabulary, it becomes possible to see how these constraints reshape academic life not only administratively but also ethically, relationally, and politically.

The findings reveal that academics experience the Third Mission primarily through activities that resemble what Arendt terms *labour*: repetitive cycles of reporting, documenting, and categorising that absorb time yet produce little lasting recognition. This experience resonates with the extensive literature on the intensification of academic workload under performance-based governance, but the Arendtian frame clarifies why these pressures feel so consuming. Labour in this sense is not only abundant; it is structurally inevitable. The Third Mission's evaluative machinery must be continually fed, and the burden falls most heavily on those with the least institutional power, a pattern that mirrors longstanding evidence on gendered and hierarchical inequities in academic service work.

In parallel, the study shows that universities construct their Third Mission reality through the production and circulation of evaluative artefacts, templates, portals, categories, classifications, which function as what Arendt calls *work*: durable objects that shape the institutional world. These artefacts stabilise a narrow vision of societal engagement, privileging activities that can be easily counted, compared, and made visible. Prior scholarship has shown how metrics shape behaviour and distort value, yet the Arendtian framing reveals a deeper process of world-making: the Third Mission becomes real only to the extent that it fits into preexisting categorisations. Relational, cultural, or community-based engagements struggle to appear within such a world, regardless of their societal significance.

The implications for what Arendt describes as *action* are substantial. Although academics articulate a vision of engagement grounded in relationality, plurality, and responsiveness to local needs, the institutional conditions for such action remain fragile. The Third Mission is often framed as a space for collaboration and participation, yet evaluation systems reward high-visibility initiatives, impressive numbers, and fast results. These dynamics echo research showing that organisations routinely substitute symbolic for substantive participation. Through Arendt's lens, such findings suggest that the Third Mission risks losing its political and relational essence: instead of enabling spaces where academics and communities can appear to one another as plural actors, universities increasingly stage engagement as a performance oriented toward auditability.

Despite these pressures, the findings demonstrate that academic judgment has not disappeared. Many academics continue to articulate ethical orientations that stand in clear opposition to institutional priorities, grounded in long-term responsibility, relational trust, and attentiveness to societal consequences rather than performative outputs. Yet this persistence of judgment exposes a stark institutional contradiction. Ethical reflection survives, but it is structurally sidelined: it circulates at the level of individual reasoning while remaining largely excluded from organisational decision-making. The result is not a lack of judgment, but its containment. Within the Third Mission, societal engagement is actively encouraged as discourse, while the conditions necessary for ethical judgment to inform priorities, resource allocation, or evaluation are systematically absent. This split between thinking and acting, central to Arendt, highlights a broader paradox within the Third Mission. Universities encourage societal engagement rhetorically but offer few organisational pathways through which ethical reflection can meaningfully influence institutional priorities. Universities thus rely on the moral labour of academics while neutralising its institutional force, producing a form of engagement that is ethically articulate yet organisationally inert.

Compounding these tensions is the emergence of *thoughtlessness*, Arendt's term for unreflective adherence to rules and routines. The study shows how procedural compliance becomes normalised: autonomy is formally affirmed but substantively steered; consultation processes invite participation but rarely accommodate dissent; academics adapt their work to satisfy reporting requirements even when doing so distorts meaningful engagement. While performance-based governance has long been criticised for generating ritualism and symbolic behaviour, the Arendtian perspective reveals the potential erosion of critical thinking itself. When visibility replaces value, and compliance replaces judgment, the university risks becoming an institution that performs societal relevance rather than enacting it.

What emerges, therefore, is not merely a misalignment between metrics and meaning, but a more subtle political effect. The ambiguity that characterises Third Mission evaluation does not simply reflect compromise between competing interests; it actively enables the replacement of action and judgment with labour and compliance. By keeping the boundaries of what "counts" perpetually unstable, evaluative regimes sustain a condition in which academics must continuously adapt, translate, and justify their engagement, while rarely being able to contest the terms on which it is assessed. In this sense, ambiguity functions as a low-intensity technology of power: it does not silence critique but renders it institutionally inconsequential. Thinking persists, but it is disconnected from decision-making; engagement continues, but increasingly in forms shaped by visibility rather than responsibility. The risk is not the disappearance of the Third Mission, but its hollowing out, its transformation into a procedural performance that mimics public value while progressively undermining the conditions for genuine public action.

These insights collectively point to the need for a more fundamental rethinking of how universities design and govern the Third Mission. This study contributes to existing literature by offering a nuanced theoretical account of how evaluative infrastructures reshape academic agency. By incorporating Arendt's distinctions, it reveals the multi-layered dynamics through which metrics and managerial tools influence not only what academics do, but how they understand their responsibilities, their relationships with communities, and their capacity to appear as reflective, plural, and politically engaged actors. The Third Mission thus emerges as a site where broader tensions in contemporary university governance are crystallised: between visibility and meaning, between compliance and reflection, between managerial order and public responsibility.

The study's insights should, however, be interpreted with acknowledgement of its limitations. It focuses on a single national context, shaped by Italy's distinctive governance structures, evaluation frameworks, and funding mechanisms. While many dynamics observed here echo those found in other higher education systems, comparative research would help clarify how national institutional arrangements shape the interplay between evaluation and engagement. Moreover, the study draws primarily on the perspectives of academics. Arendt's emphasis on plurality suggests the importance of including the voices of community partners, NGOs, cultural institutions, and local actors who co-produce engagement but navigate institutional pressures from the outside. Additional methodological approaches, such as ethnographic observation of evaluative and participatory processes, could further illuminate how thoughtlessness, judgment, and action unfold in everyday organisational practice. Yet, an additional crucial aspect refers to a disciplinary specificity. While the present study deliberately adopts a cross-disciplinary perspective to capture how Third Mission metric regimes are experienced across the contemporary university, future studies could profitably examine how these dynamics unfold within a single disciplinary community. In particular, focusing on accounting academics across multiple universities would enable researchers to explore how disciplinary cultures, professional identities, and accounting-specific institutional logics shape responses to evaluative regimes. Such work would complement the present study by sacrificing disciplinary breadth for

analytical depth and could adopt a comparative cross-institutional design, following the approach developed by Spanò et al. (2024).

These limitations also point toward practical and policy implications. If universities are serious about supporting societal engagement, they must reconsider the dominance of indicators that privilege visibility over value. Metrics may be necessary, but they cannot constitute the primary language through which engagement is understood. Policymakers and university leaders should create evaluative frameworks that recognise the temporal, relational, and unpredictable character of action, rather than compelling academics to retrofit engagement into narrow templates. This might involve shifting emphasis from counting outputs to documenting relational processes, providing structural support for long-term collaborations, reducing reliance on performance-based funding for engagement, and designing participatory governance arrangements that incorporate meaningful academic and community input. Crucially, institutions must cultivate conditions that allow judgment to influence decision-making, something that cannot be achieved through formal consultation alone but requires genuine openness to plurality, disagreement, and ethical deliberation.

Ultimately, the study suggests that the Third Mission can only fulfil its societal promise if universities protect and nurture the capacities that Arendt identifies as constitutive of public life: the ability to act with others in conditions of plurality, to reflect critically on the meaning of one's responsibilities, and to resist the drift toward bureaucratic thoughtlessness. Without such institutional commitments, the Third Mission risks becoming an activity that is measurable but not meaningful, a performance of engagement rather than its enactment.

Beyond its implications for universities, this study speaks to a broader challenge for accounting as a field of practice and inquiry. The findings suggest that contemporary forms of accounting, particularly when organised around metrics, indicators, and reporting templates, risk being reduced to technical exercises centred on measurement, calculation, and procedural compliance. As Carnegie et al. (2021a, 2021b, 2024, 2025) have argued, however, accounting cannot be understood solely as a technical practice. It is also a social and moral practice that shapes what is recognised as valuable, which activities are prioritised, and how organisations relate to the wider world. Rather than asking only "*How do we account?*", the findings invite a more fundamental question: "*What kinds of realities does accounting help to produce?*" If accounting systems shape organisational priorities rather than merely represent them, then attention must also turn to "*What kinds of organisations do our accounting systems encourage?*" and "*What forms of value do they make visible or leave invisible?*" From this perspective, the dynamics observed in this study point to a broader risk: accounting systems that are narrowly oriented towards standardisation, comparability, and visibility may inadvertently privilege procedural alignment over reflective judgment. The challenge, therefore, is not to abandon measurement, but to ensure that accounting continues to support reflection, responsibility, and meaningful engagement alongside procedural accountability. Otherwise, universities risk becoming increasingly accountable in procedural terms while progressively detached from the meaning and purpose that the Third Mission is intended to serve.

This calls for a rethinking of accounting not merely as a technical practice, but as a multidimensional activity that is inherently social, and moral, and that carries responsibility for how organisations engage with complex societal challenges. In this sense, the study can be read as a broader invitation to reconsider the role of accounting in addressing "*big questions*" and "*wicked problems*," moving beyond narrow metric-driven approaches toward forms of evaluation that sustain judgment, plurality, and meaningful action. If left unexamined, this trajectory risks entrenching forms of accountability that are increasingly precise in measurement yet increasingly limited in their capacity to capture what truly matters.

In closing this paper, we warn readers that, as Arendt (1971, p. 438) reminds us, “*the sad truth is that most evil is done by people who never make up their minds to be good or evil.*” Read in the current context of contemporary universities (but not only!), this is not an accusation but a warning: when evaluative routines absorb attention and judgment is suspended, the erosion of public responsibility can occur without intent, conviction, or debate. The effectiveness of public responsibility, therefore, may be adversely impacted.

References

- Aboubichr, B., & Conway, N. (2023). The gaming of performance management systems in British universities. *Human relations*, 76(4), 602-628.
- Alvesson, M. (2003). Methodology for close up studies—Struggling with closeness and closure. *Higher Education*, 46(2), 167–193. <https://doi.org/10.1023/A:1024716513774>
- Alvesson, M., & Kärreman, D. (2007). Constructing mystery: Empirical matters in theory development. *Academy of management review*, 32(4), 1265-1281.
- Andersen, L. B., Boesen, A., & Pedersen, L. H. (2016). Performance in public organizations: Clarifying the conceptual space. *Public Administration Review*, 76(6), 852–862. <https://doi.org/10.1111/puar.12578>
- Arendt, H. (1958). *The human condition*. University of Chicago Press.
- Arendt, H. (1963). *Eichmann in Jerusalem: A report on the banality of evil*. Viking Press.
- Arendt, H. (1971). Thinking and moral considerations: A lecture. *Social research*, 417-446.
- Arendt, H. (1978). *The life of the mind*. Harcourt.
- Argento, D., & van Helden, J. (2023). Are public sector accounting researchers going through an identity shift due to the increasing importance of journal rankings? *Critical Perspectives on Accounting*, 96, 102537.
- Benneworth, P., & Jongbloed, B. W. (2010). Who matters to universities? A stakeholder perspective on humanities, arts and social sciences valorization. *Higher Education*, 59(5), 567–588.
- Bracci, E., Saliterer, I., Sicilia, M., & Steccolini, I. (2021). Accounting for (public) value(s): Reconsidering publicness in accounting research and practice. *Accounting, Auditing & Accountability Journal*, 34(7), 1513–1526.
- Carnegie, G. D. (2022). Global university rankings: The macro-micro contradiction in public university management. *Accounting and Management Review* | *Revista de Contabilidade e Gestão*, 26(1).
- Carnegie, G. D. (2023). Global university rankings and the metrics maze. *Accounting, Auditing & Accountability Journal*, 36(1), 430-433.
- Carnegie, G., Parker, L., & Tsahuridu, E. (2021a). It's 2020: What is accounting today? *Australian Accounting Review*, 31(1), 65–73.
- Carnegie, G., Parker, L., & Tsahuridu, E. (2021b, April 6). *Redefining accounting for tomorrow?* IFAC Knowledge Gateway. <https://www.ifac.org/knowledge-gateway/discussion/redefining-accounting-tomorrow>
- Carnegie, G. D., Gomes, D., Parker, L. D., McBride, K., & Tsahuridu, E. (2024). How accounting can shape a better world: Framework, analysis and research agenda. *Meditari Accountancy Research*, 32(5), 1529–1555. <https://doi.org/10.1108/MEDAR-08-2023-2108>
- Carnegie, G. D., Parker, L., & Twyford, E. (2025, February 10). *The case for redefining accounting as a multidimensional practice: Shaping a better world*. *The CPA Journal*. <https://www.cpapjournal.com/2025/02/10/the-case-for-redefining-accounting-as-a-multidimensional-practice/>
- Dar, M., Cordery, C., De Loo, I., & Manochin, M. (2025). An autoethnographic reflection on accounting rituals in a turbulent environment. *Qualitative Research in Accounting & Management*, 22(1), 104-128.
- Deem, R., & Brehony, K. J. (2005). Management as ideology: The case of 'new managerialism' in higher education. *Oxford Review of Education*, 31(2), 217–235. <https://doi.org/10.1080/03054980500117827>
- Drujon d'Astros, C., Gaudy, C., & Strauch, M. (2025). The ethnographer and the paradox of emotional closeness. An autoethnography of shame in accounting research. *Qualitative Research in Accounting & Management*, 22(1), 65-86.
- Gebreiter, F. (2022). A profession in peril? University corporatization, performance measurement and the sustainability of accounting academia. *Critical Perspectives on Accounting*, 87, 102292.

- Gendron, Y., & Rodrigue, M. (2021). On the centrality of peripheral research and the dangers of tight boundary gatekeeping. *Critical Perspectives on Accounting*, 76, 102076.
- Haynes, K. (2017). Autoethnography in accounting research. In *The Routledge companion to qualitative accounting research methods* (pp. 215-230).
- Haynes, K. (2013). Sexuality and sexual symbolism as processes of gendered identity formation: An autoethnography of an accounting firm. *Accounting, Auditing & Accountability Journal*, 26(3), 374-398.
- Haynes, K. (2011). Tensions in (re)presenting the self in reflexive autoethnographical research. *Qualitative Research in Organizations and Management*, 6(2), 134-149.
- Holt, R. (2020). Hannah Arendt and the raising of conscience in business schools. *Academy of Management Learning & Education*, 19(4), 584-599.
- Hopwood, A. G. (1990). Accounting and organisation change. *Accounting, Auditing & Accountability Journal*, 3(1).
- Kallio, K.-M., Kallio, T. J., & Grossi, G. (2017a). Performance measurement in universities: Ambiguities in the use of quality versus quantity in performance indicators. *Public Money & Management*, 37(4), 293-300.
- Kallio, K.-M., Kallio, T. J., Tienari, J., & Hyvönen, T. (2017b). Ethos at stake: Performance management and academic work in universities. *Human Relations*, 70(2), 239-262. <https://doi.org/10.1177/0018726716652427>
- Loi, M. & Di Guardo, M.C. (2015). The third mission of universities: An investigation of the espoused values. *Science and Public Policy*, 42(6), 855-870.
- Manes Rossi, F., Spanò, R., Martin-Sardesai, A., & Guthrie, J. (2022). Performance management change: Discourses at play in an Italian university. *Journal of Public Budgeting, Accounting & Financial Management*, 34(4), 534-556.
- Martin-Sardesai, A., Guthrie, J., & Tucker, B. P. (2020). What you see depends on where you look: Performance measurement of Australian accounting academics. *Accounting, Auditing & Accountability Journal*, 33(6), 1193-1218.
- Michelon, G. (2021). Accounting research boundaries, multiple centers and academic empathy. *Critical Perspectives on Accounting*, 76, 102204.
- Morsing, M., & Schultz, M. (2006). Corporate social responsibility communication: Stakeholder information, response and involvement strategies. *Business Ethics: A European Review*, 15(4), 323-338. <https://doi.org/10.1111/j.1467-8608.2006.00460.x>
- Parker, L. D. (2024). Public university research engagement contradictions in a commercialising higher education world. *Financial Accountability & Management*, 40(1), 16-33. <https://doi.org/10.1111/faam.12341>
- Parker, L. D., Guthrie, J., & Martin-Sardesai, A. (2024). Performance management in the Australian higher education system—a historically informed critique. *Accounting History*, 29(2), 215-235.
- Sadler, E., & Wessels, J. S. (2019). Transformation of the accounting profession: An autoethnographical reflection of identity and intersectionality. *Meditari Accountancy Research*, 27(3), 448-471.
- Sánchez-Barrioluengo, M. (2014). Articulating the ‘three-missions’ in Spanish universities. *Research Policy*, 43(10), 1760-1773.
- Sicilia, M., Steccolini, I., Mauro, S.G., Barbera, C. and Grossi, G. (2025). Half a century of public sector accounting research through bibliometric analysis: Looking back to move forward. *Financial Accountability & Management*, <https://onlinelibrary.wiley.com/doi/full/10.1111/faam.12432>
- Spanò, R., Bracci, E., Manes-Rossi, F., & Sforza, V. (2024). A fatally efficient machine. Insights into the ‘banality’ of the research evaluation exercise in Italy. *Critical Perspectives on Accounting*, 100, 102742.
- Spanò, R., Grossi, G., & Landi, G. C. (2022). Academic entrepreneurial hybrids: Accounting and accountability in the case of MegaRide. *The British Accounting Review*, 54(5), 101130.

Steccolini, I. (2019). Accounting and the post-New Public Management: Re-considering publicness in accounting research. *Accounting, Auditing & Accountability Journal*, 32(1), 255–279.

Tanima, F. A., Brown, J., & Hopper, T. (2024). Doing critical dialogic accounting and accountability research: an analytical framework and case illustration. *Accounting, Auditing & Accountability Journal*, 37(1), 1-30.

ter Bogt, H. J., & Scapens, R. W. (2012). Performance management in universities: Effects of the transition to more quantitative measurement systems. *European Accounting Review*, 21(3), 451–497. <https://doi.org/10.1080/09638180.2012.668323>

Zou, Y. (2021). Methodological insights experiencing and knowing in the field: an autoethnographic account of conducting audit fieldwork in China. *Accounting, Auditing & Accountability Journal*, 34(3), 680-698.