SUSTAINABILITY REPORTING FOR AND BY CITIES

A LONGITUDINAL ANALYSIS OF EUROPEAN AND LATIN AMERICAN PRACTICES

Ludger Niemann

TU Delft

Sustainability reporting for and by cities.

A longitudinal analysis of European and Latin American practices

Dissertation

for the purpose of obtaining the degree

of doctor at Delft University of Technology

by the authority of the Rector Magnificus prof.dr.ir. H. Bijl

Chair of the Board for Doctorates

to be defended publicly on

Thursday, 29 January 2026 at 17:30

by

Ludger NIEMANN

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof. Mr. Dr. J.A. de Bruijn	Delft University of Technology, promotor
Prof. Dr. T. Hoppe	University of Twente, promotor

Independent members:

Prof. Dr. E.M. van Bueren	Delft University of Technology
Prof. Dr. Ir. C.A. Ramírez Ramírez	Delft University of Technology
Prof. Dr. J. Edelenbos	Erasmus University Rotterdam
Prof. Dr. W.M. de Jong	Erasmus University Rotterdam
Prof. Dr. R. Torenvlied	University of Twente
Prof. Dr.. M.J.G. van Eeten	Delft University of Technology, reserve member

ISBN: 978-94-6518-221-6

DOI: A DOI will be assigned upon final deposit in the TU Delft Repository

An electronic version is available at http://repository.tudelft.nl

Cover design: Vladimir Victor Ceric

Printed by: Proefschrifterij, Soest

Preface and acknowledgements

Working with Ecuadorian municipalities in the early 2010s to develop monitoring and indicator systems, I became increasingly curious about how such practices were organised elsewhere. True to the Dutch adage *beter goed gejat dan slecht bedacht* (better well stolen than badly invented), I began looking for examples of promising approaches in other contexts.

An online search led me to a publication from the University of Twente regarding the "Local Sustainability Meter." Intrigued, I contacted the authors, Thomas Hoppe and Frans Coenen, and we arranged a pleasant video call. When I asked whether they were aware of similar research in other countries, the response was straightforward: not really—but why not take it on myself? That friendly challenge marked the beginning of a long academic journey.

This project unfolded in parallel with my relocation to the Netherlands in 2014, gradually crystallizing into a comparative investigation of sustainability reporting practices in Europe and Latin America. Along the way, I learned extensively from a body of knowledge immediately relevant to policy practice. I also realised that I initially held a somewhat naive belief in the immediate, transformative power often attributed to transparency and accountability tools. As I engaged more deeply with the academic literature—and particularly through the empirical studies presented in this dissertation—I developed a more nuanced perspective. While these instruments can be transformative, their effects are highly contingent on political, organisational, and contextual conditions. This evolution has strengthened my conviction that systematic evaluation remains a crucial mechanism for improving real-world practices.

I would like to express my profound gratitude to those who supported this work. My deepest thanks go to Thomas Hoppe and Hans de Bruijn. Thomas, whose role evolved from daily supervisor to *promotor*, provided exceptional reliability and constructive feedback from the very beginning of this project. Hans, who served as *promotor* during my time at TU Delft, provided highly insightful and patient guidance that was instrumental in shaping this dissertation. I also thank Hans Bressers and Frans Coenen for their support during the earliest phases of this research, and the members of the doctoral committee for their thoughtful and valuable feedback on this work.

This research would not have been possible without the many civil society activists from the *Red Latinoamericana de Ciudades Justas y Sustentables* and the local government officials in Latin America and Europe who generously shared their time, insights, and experiences. *Mil gracias, muito obrigado, dank jullie*

wel, vielen Dank, and *grand merci.* I am also thankful to GIZ and VNG International for supporting parts of this work at the institutional level.

I extend my appreciation to Erasmus University Rotterdam, where completing a Master's in Public Administration in 2016 provided an essential bridge between my earlier training in psychology and the fields of public management. I also thank my numerous colleagues at The Hague University of Applied Sciences, particularly within the Faculty of Public Management, Law and Safety and the research group led by Martijn van der Linden, where follow-up projects are now taking shape.

Finally, I am deeply grateful to my family, including my beloved three children, and to numerous friends for their enduring support throughout the years in which this dissertation was researched and written. I also thank Melanie and Emiel for their willingness to stand by my side as *paranimfs*.

It is my sincere hope that the collective work of civil society activists, public administrators, businesses, and academic researchers around the world will continue to strengthen urban sustainability practices—a task that remains critically urgent in an unsustainable world.

Amersfoort, December 2025

Summary in English

This dissertation, titled "Sustainability Reporting for and by Cities: A Longitudinal Analysis of European and Latin American Practices," explores how sustainability reporting initiatives can contribute to better urban governance through their influence on decision-making, learning, and public awareness. This research addresses a significant gap in the sustainability governance literature: although the use of indicators and reporting mechanisms has become widespread, particularly in connection with international frameworks such as the Sustainable Development Goals (SDGs), there is limited longitudinal and comparative research on how reporting initiatives function in practice—particularly at the local level and in non-Western contexts.

To address this knowledge gap, this dissertation pursues the following main research question: *How can sustainability reporting initiatives benefit decision-making, learning, and public awareness in the divergent contexts of European and Latin American cities?*

This research challenge is increasingly relevant. As sustainability reporting obligations grow—spurred by national regulations, global reporting standards (e.g., ISO), and SDG-related monitoring frameworks—so too does the expectation that more data will automatically lead to better governance. Some policymakers and civil society actors assume that the act of reporting itself will drive policy change or public accountability. However, as this dissertation shows, such assumptions can be naïve. Real-world initiatives reveal complex dynamics shaped by context, design choices, and stakeholder engagement—and are not uniformly effective.

To explore these dynamics, the dissertation undertakes a longitudinal, comparative analysis of pioneering initiatives in Europe and Latin America. It applies a theory-informed analytical framework developed iteratively through four empirical studies, which together illustrate a spectrum of "information-based governance" ranging from soft monitoring to hard accountability:: (1) municipal reporting in the European cities of Amsterdam, Basel, Dublin, Freiburg, Nuremberg, and Zurich; (2) civil society-led indicator initiatives in 49 Latin American cities; (3) a national, policy-oriented benchmarking initiative in Dutch municipalities; and (4) performance reporting obligations in the Latin American cities of Bogotá, Córdoba, Guadalajara, and São Paulo.

Methodologically, the dissertation employs a mixed methods, longitudinal multi-case study design, combining document analysis, expert interviews, and embedded case comparisons. Case selection focused on 'frontrunner' cities and

reporting systems, enabling the study of long-term developments in institutional design, stakeholder dynamics, and observed effects. The analytical framework distinguishes between three key dimensions: (i) context (e.g., political regime, administrative capacity, civil society strength); (ii) design (e.g., objectives, governance arrangements, indicator choices, dissemination strategies); and (iii) effects (e.g., instrumental, conceptual, and political-symbolic use and influence). This framework was progressively refined and ultimately culminated in a matrix of design strategies and an evaluation model linking context and design to observed outcomes.

The findings demonstrate that there is no universal model of sustainability reporting that guarantees positive outcomes. Rather, context-specific design configurations matter deeply and often act to counter-balance local political features (the "paradox of context"). In consensual, high-capacity political systems (e.g., the Netherlands, Switzerland), collaborative and benchmarking-oriented models tend to foster organisational learning and occasionally influence policy uptake by introducing artificial competition. In more adversarial or unstable political contexts (e.g., Brazil or Mexico), civil society-led or legally mandated tools can enhance transparency by providing a neutral "safe harbour" for data, but often face political resistance, data limitations, and credibility challenges.

Across all cases analysed, the most consistently observed benefit was organisational learning, particularly in the early phases of implementation. Effects such as policy influence, public awareness, and enhanced legitimacy varied more substantially and were shaped by the design, timing, and inclusiveness of each initiative. Importantly, this dissertation identifies several risks and limitations: unrealistic expectations often lead to disillusionment; rigid accountability mechanisms may induce strategic ‘gaming’ of indicators; and superficial reporting formats can undermine the credibility of conclusions. Furthermore, the presence of competing reporting tools and the local political climate shape the visibility, uptake, and institutional support for any given initiative.

The empirical studies also underscore the importance of periodicity, rotating thematic focus, and adaptive strategies for maintaining stakeholder engagement and public interest. Reporting fatigue, diminishing media attention, or institutional inertia are common threats—especially when reports become routine or detached from decision-making processes.

Academically, the dissertation contributes to three main areas. First, it advances the sustainability governance literature theoretically and empirically by offering a nuanced understanding of how reporting tools function across divergent institutional settings and over time identifying the specific "ecological niche" required for initiatives to survive. Second, it offers methodological innovation by

combining comparative and longitudinal approaches with theory-driven evaluation logic. Third, it bridges the gap between theoretical and practical work by offering a refined framework for assessing the use and influence of sustainability reporting systems.

Practically, the dissertation provides actionable guidance for policymakers and practitioners. It identifies five interlinked domains that shape the effectiveness of sustainability reporting: (1) understanding the local context; (2) clarifying purpose and anticipated effects; (3) choosing appropriate governance and organisational arrangements; (4) structuring content and communication; and (5) embedding evaluation and adaptability. Each domain is elaborated through cross-case insights and synthesised in Chapter 6, including a boxed summary for policy use.

In conclusion, sustainability reporting can be a valuable tool for urban governance—but only when its design is carefully tailored to context, its purposes are clearly articulated, and its implementation is flexible and adaptive. Rather than promoting sustainability reporting as a one-size-fits-all solution, this dissertation encourages researchers and practitioners to embrace the complexity of local contexts and focus on designing initiatives that are likely to generate real and lasting benefits. In doing so, the study offers a platform for more grounded, comparative, and theory-informed research on urban sustainability governance.

Nederlandstalige samenvatting

Deze dissertatie, getiteld *"Sustainability Reporting for and by Cities: A Longitudinal Analysis of European and Latin American Practices,"* onderzoekt hoe duurzaamheidsrapportage-initiatieven kunnen bijdragen aan beter stedelijk bestuur door hun invloed op besluitvorming, leren en publieke bewustwording. Dit promotieonderzoek adresseert een belangrijke lacune in de literatuur over duurzaamheidsbestuur: hoewel het gebruik van indicatoren en rapportage-mechanismen wereldwijd sterk is toegenomen—met name in verband met internationale raamwerken zoals de Sustainable Development Goals (SDG's)—is er weinig longitudinaal en vergelijkend onderzoek naar hoe dergelijke rapportage-initiatieven in de praktijk functioneren, vooral op lokaal niveau en in niet-Westerse contexten.

Om deze lacune te vullen, stelt de dissertatie de volgende onderzoeksvraag centraal: *Hoe kunnen duurzaamheidsrapportage-initiatieven bijdragen aan besluitvorming, leren en publieke bewustwording in de uiteenlopende contexten van Europese en Latijns-Amerikaanse steden?*

Deze onderzoeksvraag is bijzonder actueel. Naarmate verplichtingen tot duurzaamheidsrapportage toenemen—gestimuleerd door nationale regelgeving, mondiale standaarden (zoals ISO) en SDG-gerelateerde monitoringskaders—groeit ook de verwachting dat meer data automatisch zal leiden tot beter bestuur. Sommige beleidsmakers en maatschappelijke actoren gaan ervan uit dat het rapporteren op zichzelf al beleidsverandering of publieke verantwoording teweegbrengt. Dit proefschrift laat echter zien dat dergelijke aannames vaak te optimistisch zijn. Rapportage-initiatieven in de praktijk tonen complexe dynamieken, gevormd door context, ontwerpkeuzes en betrokkenheid van belanghebbenden—en ze zijn lang niet altijd even effectief.

Om deze dynamiek te onderzoeken, is in dit promotieonderzoek een longitudinale, vergelijkende analyse uitgevoerd naar toonaangevende initiatieven in Europa en Latijns-Amerika. Er wordt een theoretisch onderbouwd analytisch raamwerk toegepast dat stapsgewijs is ontwikkeld met vier empirische studies, die gezamenlijk een spectrum van 'informatiegestuurde governance' illustreren, variërend van zachte monitoring tot hard moeten afleggen van verantwoordelijkheden:: (1) gemeentelijke rapportage in de Europese steden Amsterdam, Basel, Dublin, Freiburg, Neurenberg en Zürich; (2) door het maatschappelijk middenveld geleide indicatorinitiatieven in 49 Latijns-Amerikaanse steden; (3) een nationaal benchmarking-initiatief gericht op duurzaamheidsbeleid van Nederlandse gemeenten; en (4) prestatiegerichte rapportageverplichtingen in de Latijns-Amerikaanse steden Bogotá, Córdoba, Guadalajara en São Paulo.

Methodologisch hanteert dit promotieonderzoek een "mixed methods", longitudinale case study-benadering, waarin documentanalyse, expertinterviews en vergelijkende casusanalyse worden gecombineerd. De selectie van steden en initiatieven richt zich op zogenaamde 'frontrunners', wat het mogelijk maakt om langetermijn-dynamieken te bestuderen, zoals institutionele continuïteit,

veranderende effecten in de tijd, en evoluerende betrokkenheid van stakeholders. Het analytische raamwerk onderscheidt drie kerncomponenten: (i) context (bijv. politiek regime, bestuurscapaciteit, kracht van het maatschappelijk middenveld); (ii) ontwerp (bijv. doelstellingen, governance-arrangementen, keuze van indicatoren, wijze van verspreiding); en (iii) effecten (bijv. instrumenteel, conceptueel en politiek-symbolisch gebruik en invloed). Het raamwerk is door empirisch werk verfijnd en heeft uiteindelijk geresulteerd in een matrix van ontwerpstrategieën en een evaluatiemodel dat context en ontwerp verbindt met waargenomen uitkomsten.

De bevindingen tonen aan dat er geen universeel model voor duurzaamheidsrapportage bestaat dat gegarandeerd positieve resultaten genereert. Context-specifieke ontwerpconfiguraties zijn van cruciaal belang en fungeren vaak als tegenwicht voor lokale politieke kenmerken (de 'paradox van de context'). In consensusgerichte, goed functionerende bestuursstelsels (zoals Nederland en Zwitserland) hebben collaboratieve en op benchmarking gebaseerde modellen de neiging om organisatiegericht leren te bevorderen. Zij kunnen soms ook beleidsimpact hebben door kunstmatige competitie te introduceren. In meer conflictueuze of instabiele contexten (zoals in Brazilië of Mexico) kunnen initiatieven die door het maatschappelijk middenveld worden geleid of wettelijk verplicht zijn, transparantie vergroten door een neutrale 'veilige haven' voor data te bieden, maar ze ondervinden vaak politieke weerstand, databeperkingen en geloofwaardigheidsproblemen.

In alle onderzochte gevallen was organisatiegericht leren het meest consistent waargenomen positieve effect, vooral in de beginfase van implementatie. Andere effecten—zoals beleidsbeïnvloeding, publieke bewustwording en versterkte legitimiteit—varieerden sterk per case en werden beïnvloed door ontwerp, timing en de mate van inclusiviteit. De resultaten van het promotieonderzoek wijzen bovendien op verschillende risico's en beperkingen: onrealistische verwachtingen kunnen tot teleurstelling leiden; rigide verantwoordingsmechanismen kunnen strategisch gedrag uitlokken ('gaming'); en oppervlakkige rapportagevormen kunnen afbreuk doen aan de geloofwaardigheid van bevindingen. De aanwezigheid van concurrerende rapportagesystemen en het lokale politieke klimaat beïnvloeden eveneens de zichtbaarheid, acceptatie en institutionele inbedding van elk initiatief.

De empirische studies onderstrepen ook het belang van periodiciteit, roterende thematische focus en aanpassingsvermogen om de betrokkenheid van stakeholders en publieke interesse te behouden. Rapportagemoeheid, afnemende mediabelangstelling en institutionele inertie vormen reële risico's—vooral wanneer rapporten routinematig worden of los komen te staan van beleidsprocessen.

Dit promotieonderzoek levert bijdragen aan drie belangrijke wetenschappelijke debatten. Ten eerste verrijkt het de theorievorming en empirische kennis over lokaal duurzaamheidsbestuur door een genuanceerd inzicht te bieden in hoe rapportagemethoden functioneren in uiteenlopende institutionele contexten en over langere tijd, waarbij de specifieke 'ecologische niche' wordt geïdentificeerd die nodig is voor het overleven van initiatieven. Ten tweede biedt het methodologische vernieuwing door longitudinaal en vergelijkend onderzoek te combineren

met theoriegestuurde evaluatielogica. Ten derde overbrugt het de kloof tussen theorie en praktijk door een verfijnd raamwerk aan te reiken voor het evalueren van het gebruik en de impact van rapportagesystemen.

In praktisch opzicht biedt dit proefschrift concrete handvatten voor beleidsmakers en praktijkprofessionals. Er worden vijf onderling verbonden domeinen geïdentificeerd die bepalend zijn voor de effectiviteit van duurzaamheidsrapportage: (1) inzicht in de lokale context en het diagnosticeren van 'governance-tekorten'; (2) heldere definiëring van doelen en beoogde effecten; (3) passende bestuurlijke en organisatiemodellen; (4) inhoudelijke en communicatieve keuzes; en (5) de inbedding van evaluatie en aanpassingsvermogen. Elk domein wordt uitgewerkt aan de hand van casus-overstijgende inzichten en samengevat in Hoofdstuk 6, waar ook een overzicht met beleidsadviezen te vinden is.

Tot slot kan duurzaamheidsrapportage een waardevol instrument zijn voor lokaal bestuur—maar alleen indien dit zorgvuldig ontworpen, contextueel afgestemd en flexibel uitgevoerd wordt. In plaats van te streven naar een universele aanpak, moedigt dit promotieonderzoek beleidsmakers en onderzoekers aan om de complexiteit van lokale contexten serieus te nemen en rapportage-initiatieven zodanig vorm te geven dat ze reële en blijvende meerwaarde opleveren. Zo biedt deze studie een fundament voor toekomstig vergelijkend en theoriegestuurd onderzoek naar duurzaamheidsbestuur in steden over de hele wereld.

Table of contents

List of abbreviations

CSR	Corporate social responsibility
CSRD	Corporate Social Responsibility Directive (*European Union*)
EU	European Union
ESG	Environmental, Social and Governance
GDP	Gross Domestic Product
GRI	Global Reporting Initiative
IR	Integrated Reporting
ISO	International Standardisation Organisation
LSM	Local Sustainability Meter (Dutch: *Lokale Duurzaamheidsmeter*)
NPM	New Public Management
NUA	New Urban Agenda (*United Nations*)
PDP	Participatory and Deliberative Processes
RCT	Randomised Controlled Trial
SDG	Sustainable Development Goals (*United Nations*)
SDSN	Sustainable Development Solutions Network
WEIRD	Western, educated, industrialised, rich, and democratic
WHO	World Health Organisation

List of Tables

List of Figures

1

Introduction: The challenge of making sustainability reporting useful for cities in diverse contexts

As cities navigate the complexities of climate change, social inequality, and growing urbanisation, they face increasing pressure to demonstrate progress toward sustainability goals. Across sustainability policy, public management, modern business administration, and fields like philanthropy, there is a frequent call for data-driven and evidence-based approaches, as well as for reporting to communicate trends and actions. This trend increasingly affects, and is influenced by, policymakers, activists, and researchers at the city level, where sustainability reporting has emerged as a key tool for promoting transparency and accountability.

What is sustainability reporting? According to technical definitions, it refers to the disclosure—whether voluntary, solicited, or required—of financial and non-financial information on sustainability performance to external stakeholders (Bebbington & Larrinaga, 2014; Mol et al., 2024). Broadly defined, sustainability reporting concerns information about environmental, social, economic, and governance issues. While common in the corporate sector—where large companies publish sustainability reports annually—it has also expanded to the public sector and civil society, giving rise to diverse practices.

Beyond the technical definition, it is worth noting that in English, the term 'reporting' has several closely related meanings ('Reporting', 2024). In one sense, reporting refers simply to the act of sharing data—such as when a municipality publishes sustainability indicators online. In a narrower sense, it involves the preparation of formal documents. This dissertation focuses primarily on the latter: structured 'urban sustainability reports' produced by local governments or other actors at regular intervals.

The proliferation of such reporting naturally raises the question whether—and how—the presumed benefits of transparency and accountability are actually realised in diverse urban contexts. This requires evaluative and comparative research.

Some overviews and case studies suggest that many initiatives bring about positive changes whereas others struggle to achieve their intended effects. Some sustainability reports never leave a mark; others generate initial interest but gradually lose relevance. Understanding both intangible effects (such as transparency) and more tangible benefits for producers and users requires systematic research and suitable analytical models. Likewise, the potential downsides—including strategic behaviour, information overload, or superficiality—warrant careful investigation.

In academic literature, much attention has been paid to the content of sustainability reports, especially since their rise in the 1990s in industrialised countries. Effects have been addressed to only a limited extent—some studies assess intended benefits, while others focus on risks such as 'greenwashing', where sustainability is exaggerated for strategic gain. The question of who is best placed to

monitor and report on urban sustainability is also underexplored. Should this task fall to governments, civil society, or collaborative partnerships? Presumably, different arrangements lead to different outcomes.

Additionally, the variables of time and place constitute important research gaps. Many studies take a snapshot approach, overlooking longitudinal dynamics that influence long-term effectiveness. Further, few studies have asked why similar reporting arrangements succeed in one city but fail in another, especially across countries or continents. The existing literature on sustainability reporting is geographically imbalanced. Most empirical studies focus on practices in high-income, industrialised countries—particularly in Europe, North America, and parts of East Asia. In contrast, there is a marked scarcity of comparative or longitudinal research on reporting practices in cities across the Global South, including Latin America, Africa, and much of Asia. This publication bias limits theoretical development and policy relevance by excluding diverse political, socio-economic, and institutional contexts where sustainability reporting is also taking root. Addressing this gap is essential to broadening the global evidence base and understanding how different reporting models function beyond the Global North.

Thus, essential knowledge gaps remain in three key areas: production processes, contextual variation across space and time, and effects. Addressing these requires systematic, theory-informed, longitudinal, and comparative research—still scarce in current literature. This dissertation seeks to understand how sustainability reporting in cities can generate lasting positive effects in diverse socio-political contexts. The question acknowledges that many reports fall short, that some lose relevance over time, and that design, context, and implementation matter deeply. The goal of this dissertation is to generate theoretical insights based on empirical case studies. Its core contribution lies in evaluating the characteristics, design, and effects of city-level sustainability reporting, with a focus on Europe and Latin America. Capturing these dynamics requires the development of frameworks and methods capable of linking design and context to outcomes over time.

One final point deserves mention: Why does this research matter? What is at stake if an urban sustainability report succeeds or fails? While such reports may seem innocuous—mere pages of text, indicators, or dashboards—they can shape public debates and policymaking. Poorly designed reports not only waste financial resources—often tens of thousands of euros or dollars—but may also mislead the public or fail to support sound policy decisions. In this sense, knowledge on when and why reporting works is essential. This study therefore speaks to ongoing policy discussions. As reporting obligations and frameworks—linked, for instance, to the Sustainable Development Goals (SDGs)—continue to evolve, many policymakers wonder how to scale urban reporting from frontrunners and what types of frameworks to promote or mandate.

The remainder of this introduction is structured as follows. Section 1.2 presents operational criteria and information on prevalence. Section 1.3 reviews relevant theoretical domains. Section 1.4 identifies key knowledge gaps, research questions, and shared analytical dimensions. Section 1.5 explains the structure and 'red thread' of the dissertation, including four empirical studies (Chapters 2–5) and the final synthesis, discussion, and conclusion (Chapter 6).

1.1 City-oriented sustainability reporting: Empirical and policy perspectives

What is meant by city-oriented sustainability reporting tools? To delineate the research matter, Sections 1.1.1 and 1.1.2 present an overview of initiatives with selection criteria and real-world examples, empirical data about their prevalence, and insights on key policy drivers.

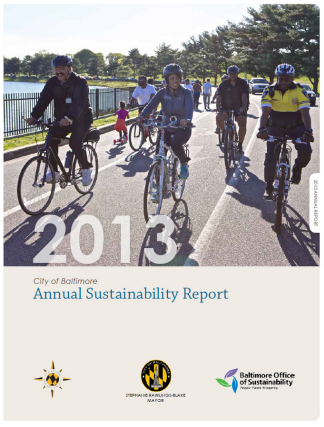
1.1.1 Operational criteria: What are reporting initiatives?

The concept of sustainability reporting has its origins in the corporate sector of the 1970s (Rimmel et al., 2024a). In the late 1980s, chemical companies with reputation problems and some environmental frontrunners started to issue environmental reports; many then broadened the scope and integrated environmental and social issues to what became 'corporate social responsibility' (CSR) reports and eventually 'sustainability reports' (Larrinaga & Bebbington, 2021). This practice is now commonplace among publicly listed corporations. More recent is the trend of merging (annual) sustainability reports with (annual) financial reports into so-called 'integrated reports'; if this trend persists, 'sustainability reporting' as a distinct activity may diminish or even disappear. In the European Union, the Corporate Sustainability Reporting Directive (CSRD) came into force in 2024 and obliges all large companies to issue annual sustainability reports (as annex to financial reports). Through reporting, business companies and industries can increase their reputation but may conversely be accused of producing sustainability reports for window-dressing or 'greenwashing' when organisational activities that superficially improve sustainability are highlighted while the core business is not addressed (Gray, 2010). For some scholars, the link with the corporate sector is so dominant that some publications do not even mention the notion of sustainability reporting applying to the public sector, let alone cities. Examples include recent articles (e.g., Baumüller et al., 2024; Runesson & Samani, 2024) and even an entire edited book on sustainability reporting (Rimmel et al., 2024b) that exclusively address the corporate sector.

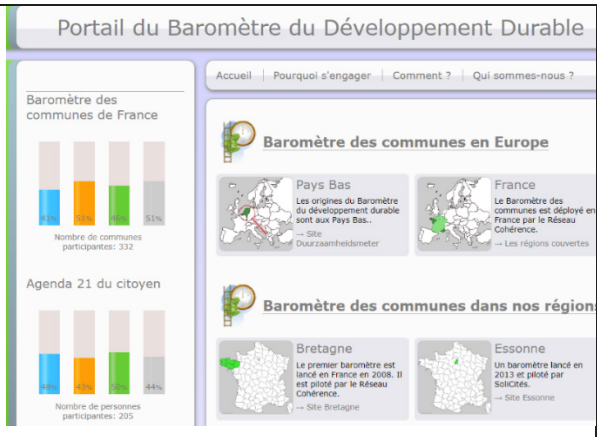
Compared with private sector organisations that tend to have clearly defined reporting obligations (typically, audited annual reports), the picture is more diffuse in the public sector. In democracies, elected governments, including municipalities, tend to take stock of their finances (and have these audited) once a year, yet usually they are not obliged to issue narrative annual reports. Instead, local

governments have larger planning and reporting cycles tied to electoral terms, as well as numerous shorter term and ad hoc reports. An example would, again, be the act of a mayor blogging about her or his administration's latest actions under the title 'sustainability report'. Furthermore, in the case of cities, the governance situation is complex for several reasons. One is that some public services (such as spatial planning or policing) are usually provided by local governments yet others (e.g., in education or health) by other layers such as national authorities. At the same time, numerous nongovernmental actors are also involved and 'make the city'. Consequently, local governments are never in charge of everything occurring in a city, nor are they necessarily the only organisations to legitimately issue city-oriented sustainability reports. In several world regions, local civil society organisations have started to compile and disclose data and issue 'sustainability reports' based on a city's indicators and trends and the actions of the city government as well as other governmental and nongovernmental actors.

'Sustainability reporting' thus appears to be an umbrella term used for diverse activities of disclosure. To illustrate this spectrum, real-world examples may be useful. Figure 1 presents the cover page of a sustainability report issued by the US city of Baltimore. This particular edition (2013) depicts Baltimore's mayor at the time and thus indicates the potential for a report to be used for political marketing purposes. The right panel of Figure 1.1 exhibits parts of a website (2024) run by a French non-profit, civil society organisation that compares and reports on the sustainability policies of various French municipalities. In this case, the organisation addressed by the report is thus different from the one issuing the report.



www.baltimoresustainability.org



www.barometredudeveloppementdurable.org

Figure 1.1 *City-oriented sustainability reporting: Real-world examples*

The example of the French, multi-city 'barometer' (which incidentally was modelled after a Dutch initiative; see Chapter 5) overlaps with benchmarking projects

and so-called 'urban observatories'. Sometimes, essentially similar activities use different labels. The opposite phenomenon involves different activities using a similar label. Both lead to blurred terminology. In their recent overview, Spada and Paulson (2023) suggested that one cannot compile coherent global maps or inventories of participatory and deliberative processes due to the 'tower of Babel' of overlapping labels.

Thus far, one can conclude that under the umbrella term of sustainability reporting, there exists a wide spectrum of practices and related activities. In light of much conceptual fluidity and no straightforward categorisation of 'sustainability reporting systems' as an empirical phenomenon, certain operational criteria that build on existing typologies are necessary to begin with (Grönholm & Berrini, 2014). This dissertation sets out to examine city-oriented sustainability reporting initiatives that meet the criteria summarised in Table 1.1. Initiatives meeting the five inclusion criteria in the left-hand column are considered examples of 'city-oriented sustainability reporting' whereas those in the right-hand column are not.

Table 1.1 *City-oriented sustainability reporting: Criteria*

City-oriented sustainability reporting	
Inclusion criteria	**Not included in this research project**
Sustainability is addressed in multiple domains	Reporting on specific domains such as solely environmental or social issues
The reporting addresses relevant trends and actions taken (and planned) by local government authorities and possibly other relevant actors	Initiatives that only release sustainability indicators or just focus on planned strategies or project activities without taking stock of the current situation
The scope encompasses action and outcomes at the city or metropolitan level	Initiatives at larger geographical scales (i.e., national, provincial level) and smaller ones (i.e., neighbourhoods)
The reporting summarises trends and actions at periodic intervals in a document	One-time assessments or city contests or certifications that are non-continuous, as well as real-time dashboards
The report is released and actively disseminated to the public	Planning and reporting systems only used by public, governmental organisations

Sustainability reporting lacks a clear-cut definition due to fluid conceptual boundaries and the proliferation of monitoring and reporting activities across overlapping policy domains. Cases in point are indicator projects relating to

specific policy sectors such as the environment as well as more generic tools addressing social accountability, citizen participation and 'open data' / government transparency. In terms of domains of tools, city-oriented sustainability reporting can arguably be viewed as a type of 'transparency and accountability initiative' that combines information on actions and achievements regarding sustainability (cf., Alcaraz-Quiles et al., 2014). From this perspective, closely related tools include community indicator projects – especially strong in North America (cf. Wray et al., 2017) – and urban observatories. In the spectrum of similar initiatives that focus on (governmental) actions, one can place participatory budgeting (cf., Manes-Rossi et al., 2023) and civil society projects aimed at tracking government expenditure. In the spectrum of similar initiatives that focus on sustainability outcomes, one can find many activities, including sustainability assessments, indicator dashboards and well-being surveys. The latter are population or household surveys on perceptions of well-being and quality of life. Figure 1.2 positions these tools along a continuum from participatory action to outcome-focused metrics, illustrating the range and hybrid nature of sustainability initiatives. The figure illustrates a sample of tools to help delineate the subject matter and does not claim to represent an exhaustive mapping.

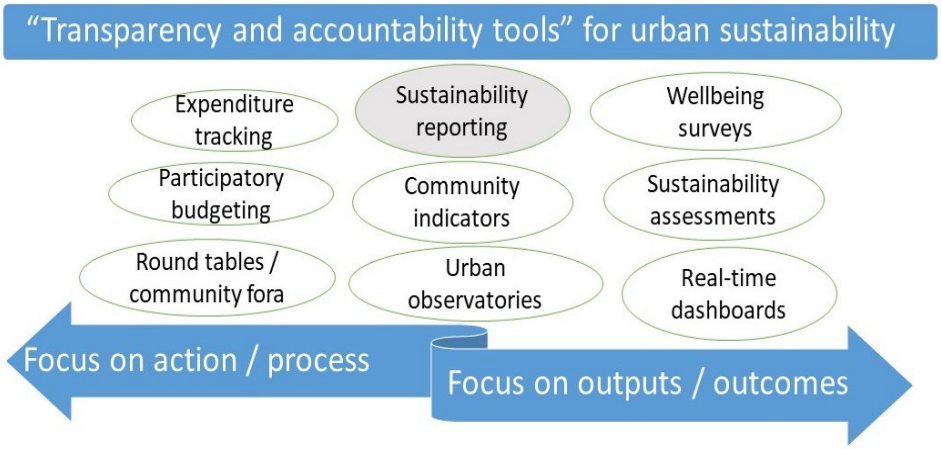


Figure 1.2 *Spectrum of tools associated with urban sustainability reporting*

To clarify the scope of this dissertation, the following **definition** of *city-oriented sustainability reporting* is proposed, grounded in the operational criteria discussed above:

> **City-oriented sustainability reporting** refers to the periodic, public dissemination of multi-domain information on sustainability trends, political commitments and actions at the city or metropolitan level. It

encompasses reports that reflect both current conditions and planned initiatives, are issued by or with the involvement of local government or civil society actors, and are intended to inform the public, support accountability, and guide local decision-making.

1.1.2 Prevalence: How common are such reports?

Since sustainability reporting remains largely voluntary in the public sector, there are no comprehensive or systematic registries of reports (Bell & Morse, 2018a). Instead, reporting trends can only be computed by researchers assessing a sample of institutional websites or by studying voluntary registries of reports that others have compiled. As indicated in the previous section, assessing the prevalence of sustainability reports is also complicated by the fact that there are no universally accepted definitions of what constitutes a sustainability report. In cities, sustainability-related information can appear in various documents, and those that meet the criteria for clear-cut sustainability reports may carry idiosyncratic titles such as "City X Progress Account".

Since the early 1990s, numerous city-focused initiatives have emerged within the so-called 'community indicators' movement. Community indicator projects typically select locally relevant sustainability indicators and publicly share data trends and sometimes relevant actions by public and private actors on websites (e.g., 'dashboards'), and in (periodic) reports addressing decision-makers and the public. Such bottom-up initiatives show considerable variations in scope – many urban monitoring and reporting initiatives are broad, considering several policy domains; others focus on environmental aspects or subjective wellbeing, often framed as 'quality of life'. Some pioneering initiatives such as 'Sustainable Seattle; changed their focus over time from broad sustainability to 'happiness' (Holden, 2013).

According to Moreno and Pires (2017), hundreds of such projects were generated in the 1990s and 2000s by a variety of actors, including civil society organisations, local governments, international philanthropic foundations, and stakeholder alliances. A 'Compendium of Sustainable Development Indicator Initiatives' launched in 2004 included at some stage over 800 initiatives in numerous countries. Some of these have ceased to exist, while others have been created more recently. Bearing in mind the challenge of unreliable registries, there is evidence that many urban indicator projects were short-lived (Moreno Pires et al., 2017). Despite their proliferation, most community indicator projects do not engage in any systematic evaluation of their work (Sirgy, 2022), and few have had discernible effects on decision-making (Hayden & Wilson, 2016). What has also not been systematically documented is the internal governance of indicator initiatives – one study examining 82 indicator projects concluded that those with a broad and cross-cutting alliance of stakeholders had the most 'staying power' (Barrington-

Leigh, 2017, p. 25); however, the sample was not necessarily representative, and this study ignored political system differences.

In addition to community-driven initiatives, a more recent trend among local governments has been the explicit use of the United Nations Sustainable Development Goals (SDGs). Although designed by and for countries, many of the framework's 17 goals and almost 200 indicators can also be related to cities. Since 2017, local and regional governments from all continents have issued 'Voluntary Local Reviews' (VLRs) on the SDGs. These are (written) assessments to monitor a jurisdiction's progress towards the SDGs) and can to some extent also be regarded as a sustainability reporting effort (Ortiz-Moya & Reggiani, 2023; Wiedmann & Allen, 2021). In sum, there are no exact numbers but it is safe to conclude that hundreds of cities on most continents have 'city-oriented sustainability reports' of some kind and that their numbers are going up.

1.1.3 Policy drivers: Who promotes such initiatives?

Having discussed the prevalence and diversity of reporting initiatives, the next logical question is: What motivates actors to initiate or maintain them? To date, only in one country – France – municipalities (with at least 50,000 inhabitants) are legally obliged to produce annual sustainability reports. This raises another central question: why do local governments or civil society organisations pursue sustainability reporting when it is not mandatory? The drivers and determinants of sustainability reporting have been a prevalent question in academic research.

In contrast to the extensive literature on corporate sustainability reporting, the literature on local governments and cities is much smaller and biassed towards certain (especially European) countries. A recent review of 190 research articles found evidence that higher reporting levels appear to be associated with larger city size, economic prosperity, and higher education levels (Mol et al., 2024). Explanations of motivations centre around the implicit objective of legitimacy gains, stakeholder management, and the 'signalling of virtue' (Low et al., 2023). There is also evidence of reporting being initiated simply because of managerial fashion (Biondi & Bracci, 2018).

Furthermore, city officials and local social activists who are concerned with sustainability issues have heeded calls for action from international institutions such as the United Nations. Since the 1990s, many institutions, including the UN, OECD, and EU, have called for the development of appropriate metrics to help guide sustainability policies. At (inter)national level, this included efforts by some organisations and commissions to develop and push for new indicators 'Beyond GDP' that are meant to counter the dominance of purely financial measures of prosperity (Hayden & Wilson, 2016). One much promoted 'alternative' indicator gaining certain recognition is the measurement of subjective well-being and

human 'happiness' which has become integrated into national (census) surveys in several countries (Krueger & Stone, 2014).

The pursuit of sustainability as a widely agreed-upon policy aim received a major boost through the Rio Earth Summit in 1992. Importantly, for cities, this framework advanced a framework (labelled *Agenda 21*) of local community dialogue, bottom-up planning and community-driven assessment tools aimed at introducing sustainability considerations into local governance (Moreno Pires et al., 2017). Although voluntary and often viewed narrowly as an environmental initiative, Local Agenda 21 (LA21) sparked a wave of global experimentation (Fenton & Gustafsson, 2017).

After the Rio+ 20 summit in 2012, the United Nations launched in 2015 the Sustainable Development Goals (SDGs) along with a specific goal to "make cities and human settlements inclusive, safe, resilient and sustainable" (UN Sustainable Development Goal 11; UN, 2015). The SDG framework – also known under the label 'Agenda 2030' – contains calls for a 'data revolution' at the national and local levels and accountability-related calls for increased reporting by all types of institutions (cf. SDG target 12.6). The SDGs have achieved significant policy traction at the level of intergovernmental organisations and national governments across the world. Contrary to the preceding Millennium Development Goals, certain SDGs (e.g., those concerning climate change) have universal applicability. As already mentioned, a number of cities in rich and poor countries have begun to issue 'Voluntary Local Reviews' on the SDGs. Policy discussions about this phenomenon are known under the heading of 'localisation' (Klopp & Petretta, 2017; United Cities and Local Governments, 2014; Zinkernagel et al., 2018).

A related milestone is the New Urban Agenda (NUA), adopted after extensive negotiations at the UN Habitat III summit in 2016 that also constitutes a landmark commitment to urban sustainability and explicitly recognised the role of subnational entities (i.e., regional and local government organisations) in achieving sustainable development (Caprotti et al., 2017; Valencia et al., 2019). As such, the SDGs and NUA were perceived as a positive milestone by some yet the policy frameworks have also been criticised as insufficiently transformative, legally non-binding, internally inconsistent, and contradictory (Del Río Castro et al., 2021).

Whereas earlier frameworks such as LA 21 in the 1990s emphasised bottom-up approaches at the local level, more recent approaches such as the SDGs explicitly or implicitly favour standardisation (Krantz & Gustafsson, 2021). One milestone in this endeavour is the set of city indicators (ISO 37120) for the "sustainable development of communities" launched by the International Standardisation Organisation (ISO, 2014). Another driver for standardisation is the emergence of numerous 'green city' ratings, rankings, and comparisons (Acuto et al., 2021; Elgert, 2018). Some of these are promoted by inter-governmental organisations

(e.g., the EU's "Green Capital Award"), others by national governments as well as businesses (Mercer, Siemens, Arcadis) and foundations (Joss et al., 2015). Another factor that is associated with standardisation is the practice emerging among (world) cities to issue periodic sustainability reports and to integrate sustainability indicators into their general planning and reporting cycles. In less-wealthy countries, multilateral financial institutions – such as the Interamerican Development Bank (IADB) – also exert considerable influence by tying the disbursement of loans and grants to the attainment of specific targets as well as the introduction of generalised result-based management and reporting systems (Feinstein & Moreno, 2015).

Contrary to the private sector, however, where certain reporting standards (notably those of the "Global Reporting Initiative", GRI) have become the norm, sustainability reports issued by local governments still tend to follow own, locally developed formats and methods (Goswami & Lodhia, 2014). The same can be said of public sector organisations in general. In the European Union, all large 'public interest entities' are required since 2014 to disclose 'non-financial and diversity information' at regular intervals (cf. Directive 2014/95/EU), yet organisations are free in the type of formats they use for this purpose.

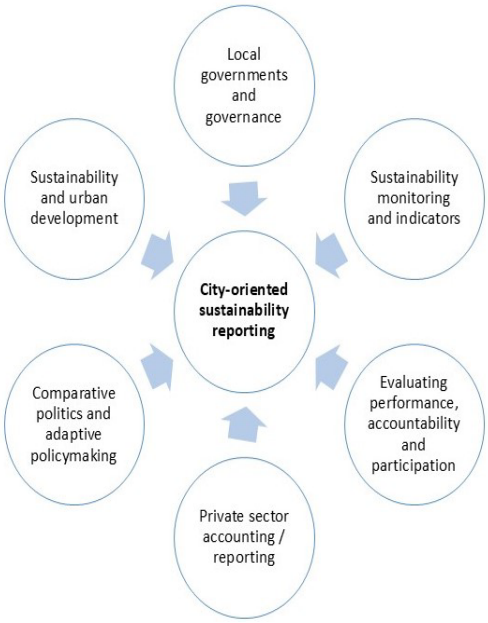
In sum, the drivers behind urban sustainability reporting are diverse. They include both local motivations—such as legitimacy and stakeholder engagement—and global institutional pressures rooted in international frameworks like the SDGs. These developments, while empirically rich, raise further questions that require theoretical framing. To explore these, the next section reviews key literatures, not by academic discipline but by topic.

1.2 Urban sustainability and governance: Theoretical perspectives

What does 'city-oriented sustainability reporting' refer to, where does it occur, and what do real-world actors, such as local governments and the United Nations, do about it? These questions have been addressed in the preceding section (1.1) designed to provide an accessible introduction to the topic. After this descriptive and largely atheoretical start, it is important to review what other researchers have found out about this phenomenon. Sustainability reporting at city level is a highly intriguing research topic because it can be – and regularly is – approached by scholars heading from various academic backgrounds. Mainstream disciplines that regularly publish about it include politics, geography, and public management but occasionally, one can also find studies on the matter in peripheral fields such as sociology, health, law, business, environmental studies, and data science. Some literature reviews have attempted to reflect on this by computing article numbers per journal (e.g., Mol et al., 2024) yet a clear-cut review by discipline is

neither feasible nor appropriate for interdisciplinary research questions and multi-author publications.

In the current case, structuring the review of the literature by topic rather than academic discipline is more appropriate. To provide an overview of essential theoretical insights on city-oriented sustainability reporting, six topics have been chosen as sub-sections (cf. Figure 1.3):



Sustainability and urban development (1.2.1); local governments and governance (1.2.2); sustainability monitoring and indicators (1.2.3); evaluating performance management, accountability and participation (1.2.4); private sector reporting and accounting (1.2.5) and, comparative politics and policy diffusion (1.2.6.). This list is certainly not exhaustive yet allows to summarise key insights from various disciplines whilst also preparing the ground for the subsequent identification of knowledge gaps.

Figure 1.3 *Academic fields offering relevant insights*

1.2.1 Sustainability and urban development

To better understand reporting processes, it is useful to first review the broader, underlying concepts of sustainability and urban development. As mentioned in Section 1.1.3, sustainability has become a mainstream, internationally accepted policy objective since the 1990s. Currently, numerous local and national governments, international organisations, civil society groups and businesses agree on its desirability. In policy frameworks as well as political science debates, the terms ‘sustainability’ and ‘sustainable development’ are often used interchangeably. The latter has been in use with since the 1980s (Moldan et al., 2012). A case in point is the ubiquitous, often-cited definition by the Brundtland report (titled “Our Common Future”) of sustainability as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment and Development, 1987, p. 43).

This interchangeable use makes explicit that sustainability is not a static state but a dynamic, future-oriented equilibrium between Earth's carrying capacity and the fulfilment of human needs. This creates inherent tensions, as numerous interests and constraints must be balanced amid changing circumstances, without certainty or consensus about the optimal equilibrium. It also deserves mentioning that the adoption of the term 'sustainable development' is contested by some scholars who repudiate it as weakening the environmental discourse or even as a case of capture by capitalist lobbies pushing the rhetoric of 'green' business (Springett & Redclift, 2017). However, it has been argued that "while scientists are endlessly discussing the details, practitioners seem to be finding their way in the forest of opinions" (Vermeulen, 2018, p. 60) since there is a broad consensus that sustainability is: i) a key international policy objective, and ii) fundamentally concerned with the core elements of "people, planet, and prosperity" (PPP).

The thinking about sustainability is of course much older than the 1990s. The concept's roots in the literature can be traced back to 1713, when the German mining administrator von Carlowitz published about *Nachhaltigkeit* (the German word for sustainability) in the context of forests affected by mines (Halla, 2021). In 1798, English economist Thomas Malthus argued that unabated growth would eventually cause all populations to succumb to famine and disease. Another publication with wide repercussions was the book *Limits to Growth* by the 'Club of Rome' in 1972. Based on computer modelling applied to population increase, agricultural production, resource depletion, industrial output, and pollution, the authors concluded that while populations grow exponentially, technology to increase resource efficiency only grows linearly, leading to unsustainable scarcity (Roseland et al., 2024).

Sustainability as a concept is thus not new. However, what changed is that von Carlowitz studied it within the localised ecosystems of specific forests, whereas people, researchers, and policymakers now discuss sustainability in terms of global level ecosystems and threats. At the same time, the local level and urban development issues have come to play a decisive role in any discussion on global sustainability. According to the UN Secretary General, "our struggle for global sustainability will be won or lost in cities" (United Nations, 2012). This is because the world is rapidly urbanising, with over 55 % of the world's population already living in cities, a figure projected to increase to about 70% by 2050 (Sharifi et al., 2024). City-dwellers also play key roles in economic networks, consumption trends, and sociocultural influences. They are connected to the rest of the world as cities are hardly ever self-sufficient but rely on their hinterland for the import of resources and the discharge of waste; as such, the term 'urban sustainability' arguably is an oxymoron, i.e., a self-contradiction (Rees, 1997). In the sustainability literature, the interplay between urban and rural systems is an important dimension in terms of so-called 'spillover effects' (Liu et al., 2018).

Cities also receive attention from (international) policymakers as local governments are seen as more innovative than national governments, making them more likely to spearhead transformative sustainability actions. They are considered to be in a better position to agree on cooperative solutions because they rely to a greater extent on face-to-face communication, which is conducive to building trust and reciprocity among stakeholders (Fuhr et al., 2018). Cities are also seen as spaces for experimentation (Kampfmann et al., 2024). On the basis of political will, behaviour change, experimentation, and technological developments, some trends towards environmental sustainability can also be discerned; prime advancements concern the increased use of renewable energy, 'water-passive' buildings, and improvements in waste collection (van der Heijden, 2014).

How to approach urban sustainability? According to Kremer et al. (2019), five main lines of thinking have been dominant in recent decades. Table 1.2 summarises the origins and main concepts of these approaches. Historical precedents include the 'Garden Cities' concept that urban planners promoted a century ago (cf. Joss, 2015), the eco-city approach, economic theories, social justice movements, and sustainability management concepts. Sustainability management has gained prominence, with some of its principles incorporated into international policy frameworks. This is the case of the Brundtland report ("Our Common Future") and the "Triple Bottom Line" concept distinguishing "people, planet, and profit" which has had significant influence on private sector reporting (cf. 1.2.5).

Table 1.2. *Origins of urban sustainability theories*

Approach	Historical provenance	Main concepts
Urban planning	• Garden Cities (early 20th century)	• New urbanism • Compact cities • Participatory planning
Greening and eco-cities	• UNESCO; COP10 (since 1970s)	• Ecosystem services • Nature-based solutions
Physical, ecological economics	• Malthusian theories (1798) • Limits to Growth (1972)	• Ecological footprint • Zero Waste / carbon pollution
Social and environmental justice	• Dumping in Dixie (1990)	• Inequitable distribution • Participation in decision-making
Sustainability management	• Our Common Future (1987) • Triple Bottom Line (1994)	• Goal setting • Continuous improvement • Indicator measurements

Underneath the surface of widespread allegiance to the global ideal of sustainability, one can thus find substantial differences in terms of concrete visions and policies. For the technologically-minded, a 'sustainable city' shows stark resemblance to a 'smart city' (de Jong et al., 2015; Schraven et al., 2021); for an environmentalist, urban sustainability is about preserving natural resources and 'nature-based solutions' (Tozer et al., 2020); from an anthropocentric perspective, urban sustainability refers to a conception of sustainability centred on the fundamental needs of individuals and 'quality of life' whereas from an "integrated perspective, urban sustainability means planning the city's future development so that its environmental impact is low and its offerings in terms of housing, work, mobility and public services remain attractive for households and businesses" (Tanguay et al., 2023, p. 104608). For some, urban sustainability also equals urban resilience, whereas others (cf. Elmqvist et al., 2019) argue that sustainability is essentially a normative and positive concept that requires "reducing unwanted resilience" and lock-ins such as urban poverty.

In summary, urban sustainability is a contested concept that is even more complex and challenging to 'implement' (Krantz & Gustafsson, 2021; Zeemering, 2018). The complexity derives from the need to coordinate diverse scales ranging vertically from the international to the very local whilst working horizontally with multiple city actors on numerous issues, including citizen participation, ecosystem services, behaviour change, technology, goals, and measurements (cf. Table 1.2). In the academic literature, all these factors and arrangements are often referred to under the heading of (local) 'sustainability governance', which the following section addresses.

1.2.2 Local governments and governance

Cities are a key unit of analysis in this project, yet they are often more diffuse constructs than commonly perceived. In particular, metropolitan areas usually have no clear boundaries but are governed by a web of local governments and other institutions, with overlapping jurisdictional scales. In the case of Mexico, for example, the city proper has about 9 million inhabitants and the metropolitan area 22 million. City or 'community' boundaries may be defined administratively, functionally (e.g., service catchment areas), or psychologically, based on residents' perceptions (Sharifi, 2016). In legal and practical terms, cities hardly ever correspond neatly to a single 'local government' and therefore not be equated in comparative studies. Instead, some researchers prefer referring to 'local authorities' which include (democratically elected) local governments as well as other governmental institutions. Furthermore, cities are generally in complex and contentious relationships with higher levels of government (Klopp & Petretta, 2017), depending on the polity or the political system in each country.

The dominant mode of urban development generates unsustainable sprawl and other externalities that city governments—typically responsible for land use and

spatial planning—must address (Correia & Roseland, 2022). In many countries, sustainability entered the city-level policy discourse through governmental planning departments, as these tend to be in charge of land use and the provision of public transport; other local governments recognised threats from global climate change and reconsidered their responsibility for carbon emissions, urban ecology, and environmental management (Zeemering, 2018). Most city governments have legal frameworks and certain policies in place about sectors that are pertinent to sustainability, such as water, housing, land use, climate change, transport, and economic development (Sharifi et al., 2024).

Another important aspect for sustainability reporting is that each local government faces choices about how to strategise and organise its work internally. Depending on legal competencies, institutional capacity, and organisational 'maturity', a municipality may adopt a broad systems perspective on sustainability. This, however, brings about complexity, which is why less experienced municipalities may find it easier to develop sustainability initiatives as separate projects running in parallel to overall management (Krantz & Gustafsson, 2021). A study of over 600 municipalities in the United States found that most adopted short-term goals in an ad hoc manner without integrated, long-term goals (Liao et al., 2020). Lastly, it has been observed that many cities (especially yet not only in developing countries) lack information about relevant trends, threats, and opportunities that would allow them to develop comprehensive sustainability strategies for local governments and other actors.

International institutions, including the UN-affiliated Sustainable Development Solutions Network (SDSN), call for a 'data revolution' that leverages technology, appropriate indicators, and enhanced data management capacities (Klopp & Petretta, 2017; SDSN, 2019). Until this has changed, the transformative potential of cities must not be overstated because local governments have limited capacities and resources, especially in the Global South. Researchers warn that without stronger local government institutions, the SDGs will not be met in most developing countries (Nunes Silva, 2020).

How can cities then be supported to work more effectively regarding urban sustainability? Some scholars and activists have derived recommendations from studies of frontrunners (cf. Fuhr et al., 2018) and call for decentralisation, capacity building, and empowerment of local governments. Other scholars focused on legal innovations, such as new taxes and tradeable land development permissions (Correia & Roseland, 2022). However, there are no universal recipes, and local contexts need to be considered. As one researcher quipped, "in developing economies, cities develop too rapidly for new legislation and regulation to be meaningful; in developed economies, cities develop too slowly for new legislation and regulation to be meaningful" (van der Heijden, 2014, p. 3).

While adequate and effective local government strategies and capacities are essential for sustainable development (as well as technological fixes and behaviour change), several researchers argue that "getting governance right" constitutes the "missing part of the puzzle" (van der Heijden, 2014, p. 13). In the context of sustainability reporting, governance is important as it provides a useful lens to study the relationship between various urban stakeholders.

What is governance? Governance, like sustainability, has gained prominence in political science and policymaking since the 1990s, with varied meanings. It can simply refer to the activity of governing ('Governance,' n.d.)—in principle even for criminal organisations. However, in policy sciences, the term is not just descriptive but also has normative connotations alluding to 'good governance'. One oft-cited definition explains governance "as a process of – more or less institutionalised – interaction between public and/or private entities aiming at the realisation of collective goals" (Lange et al., 2013, p. 406). In this definition, the allusion to a 'collective goal' makes governance a value-laden affair. For the current research project, this definition is useful since sustainability represents a collective goal and requires cooperation between i) geographical layers and levels and ii) governmental and nongovernmental actors.

Concerning geographical levels, one must bear in mind that sustainability transitions necessarily require concerted action countries, cities, neighbourhoods and citizens. Much sustainability research legitimately focuses on a particular scale, such as international policy objectives or perhaps behaviour changes and awareness among citizens or businesses. Interestingly, some scholars argue that neighbourhoods – rather than cities – constitute the most 'efficient' spatial scale for urban sustainability transitions.

Figure 1.4, adapted from Rey, Laprise and Lufkin (2022) visualises the spatial scales relevant for sustainability transitions, emphasizing the efficiency of the neighbourhood scale. For buildings and neighbourhoods, professional bodies have also developed several technically-minded sustainability frameworks (e.g., BREEAM and LEED), leading to ratings and certifications (Joss & Rydin, 2018). As the review of sustainability indicators (cf. Section 1.2.3) will show, many monitoring systems are also designed to allow for the (georeferenced) aggregation and disaggregation of data in order to visualise the interplay between scales and the contribution of, say, certain neighbourhoods to city-level trends.

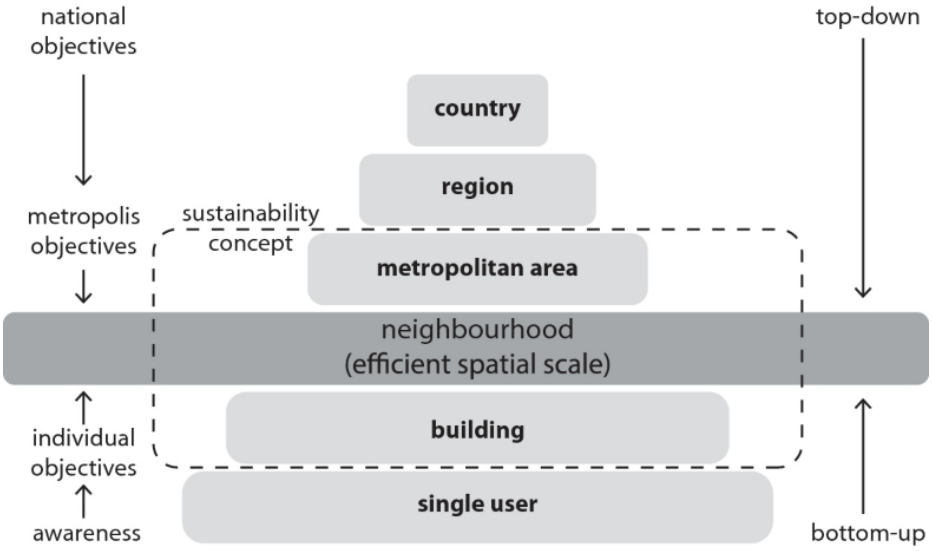


Figure 1.4. *Spatial scale for sustainability transitions as proposed by Rey et al. (2022)*

Concerning the second major aspect of urban governance – multi-stakeholder cooperation – one large research project comparing several world cities concluded that "in many local authorities, there is no shortage of ambitions in sustainability work" and that "political will, coherent governance, and strong formal partnerships between public sector, private sector and civil society actors are key ingredients" (Valencia et al., 2019, p. 20). In a similar vein, a group of researchers studying the implementation of the SDGs in world cities found that three central governance challenges are: "(i) cultivating collective action by creating inclusive decision spaces for stakeholder interaction across multiple sectors and scales; (ii) making difficult trade-offs, focusing on equity, justice and fairness; and (iii) ensuring mechanisms exist to hold societal actors to account regarding decision-making, investment, action, and outcomes" (Bowen et al., 2017, p. 91)[1]. Several dimensions in this list – in particular, stakeholder interaction, accountability, action, and outcomes – are directly relevant for the topic of city-oriented sustainability reporting.

[1] On the question of why some cities are recognised as leaders for environmental innovations, researchers identified five key drivers: i) high capacities combined with high problem pressure, ii) a well-functioning local democracy, iii) an enabling policy framework such as substantial legal competencies for the municipality, iv) socioeconomic conditions (e.g., a concerned civil-society), and v) local leadership such as a mayor with political authority (Fuhr et al., 2018).

How to pursue effective governance arrangements? As key architects of local sustainability governance, local governments or authorities can pursue various strategies at the city level. One approach to distinguish modes of action concerns 'governance styles' (i.e., hierarchical, market-based, network-based, etc.), which each have both strengths and weaknesses regarding sustainability transitions (Meuleman & Niestroy, 2015). Van der Heijden (2014) proposes three governance approaches available to local governments: (i) direct regulatory interventions (e.g., statutory regulations, taxes, subsidies); (ii) collaborative governance (networks, partnerships, covenants); and (iii) voluntary programmes and market-driven governance (benchmarking, certification, contests, sustainable procurement). The latter, market-based approach has been pushed significantly in many countries by the New Public Management paradigm of the 1990s (cf. Section 1.2.4).

According to other authors, "the classic mix of instruments in sustainability governance" (Ekardt, 2024) refers to regulations, subsidies, planning law, and informational approaches leading to "a mix of compulsion and voluntarism" (Fenton & Gustafsson, 2017, p. 129). In terms of these two mixes, sustainability reporting belongs to the category of informational governance instruments. In most countries it is still voluntary (cf. Section 1.1.3.) yet municipalities and other public sector institutions are under increasing pressure and sometimes legally obliged to be more forthcoming in disclosing indicator data and accountability-relevant information.

To sum up, for sustainable development at city level, local governments are important actors yet by far from omnipotent; depending on many contextual variables, they have some direct policy influence and a role to play in organizing local sustainability governance vertically, horizontally, and internally within the public administration. Sustainability reporting can be considered as a tool of informational governance.

1.2.3 Sustainability monitoring and indicators

As explained in the previous sections, sustainability reporting is essentially about describing and disclosing goals, actions, and outcomes. Furthermore, various urban sustainability theories rely on the continuous use of measurements (cf. Table 1.2 and Kremer et al. 2019), and policymakers call for a 'data revolution' (SDSN, 2019) to help cities. Therefore, a closer examination of theories and concepts of sustainability assessment is warranted.

For goal-setting and reporting, indicators are recurring and key components. In the field of urban sustainability assessment, indicator-based frameworks have been found to be much more prevalent than, for example, rating systems, or principle-based frameworks (Cohen, 2017). So, what exactly are indicators? A technical definition (Halla & Merino-Saum, 2022, p. 14) explains indicators as:

> multifaceted constructs ideally composed of: (1) a label; (2) a unit of measurement; (3) a definition; (4) accessible data; (5) a reference point; and (6) its anchoring in a particular conceptual framework (i.e. the categories and the concept with which it is associated). This last element differentiates an indicator from a metric, which can be seen as a 'generic' or 'raw' indicator without any concrete meaning.

In practice, however, indicators are defined in a myriad of ways. For some, an indicator is merely a source of data; for others, the term is almost equivalent to 'goal', for example when practitioners talk about "bringing some city's inequality indicators down to the national average". What most people tend to agree upon is that indicators are metrics that allow quantifying and simplifying complex social phenomena (Thomas et al., 2020), and as such they are very common tools employed to assess the performance of institutions including local governments.

According to Ramos (2019), despite the diversity of methods and tools to assess sustainability, indicator-based approaches are most commonly used since they: (i) can be selected for specific situations and needs; (ii) can be identified by stakeholders using a set of criteria; (iii) can enable a simplified representation of sustainability; and (iv) can support policy and management decisions by providing relevant performance information. Indicators are multi-use and can help in "assessing and benchmarking conditions and trends across space and time, monitoring progress toward goals and targets, informing planning and decision-making, raising awareness, encouraging political and behavioural changes, promoting public participation, and improving communication" (Klopp & Petretta, 2017, p. 94). Their appeal to decision-makers and managers stems from the rise of 'new managerialism' and 'dashboard governance' in recent decades (Thomas et al., 2020). Adherents also make the case that the current plague of 'fake news' and 'post-truth realities' can be countered by data: According to Bell and Morse, (2018b, p. 1), "we have never been so much in need of indicators to assess, in an impartial and confirmable manner, the outlines of our changing, developing, resilient, and threatened world". Naturally, the popularity of data-based approaches has historical roots too; in American cities, so-called community indicator systems based on social surveys emerged around 1910, and 'Gross Domestic Product' (GDP) as a measure of economic performance famously became popular after the Great Depression (Phillips, 2003).

In the context of performance evaluation, one frequently encounters indicator typologies. Most popular is the distinction between inputs, outputs, outcomes, and impacts, which is also known as 'results chain' or 'pipeline logic model' and represents a key tool in 'results-based management' (Lewis, 2015). Consider the case of a local government setting up and then measuring the progress of a set of activities aimed at improving local air quality. As visualised in Table 1.3 (adapted from Huovila et al. 2019), the local government in question can conceptualise and

measure its activities – such as the procurement of tree seedlings – as 'inputs', then count the number of effectively planted trees as 'outputs', and measure the (increased) amount of urban tree coverage as 'outcomes', before hopefully being able to observe improvements in air quality as 'impact'. As also explained in Table 1.3, input and process indicators are used for planning and monitoring purposes, whereas outputs can only be assessed post-implementation (i.e. after trees have been planted); the assessment of outcomes and impacts requires yet larger timespans. However, with increasing temporal and causal distance from the intervention, the challenge of establishing attribution becomes more pronounced.

Table 1.3 *Typology of indicators*

Type of indicator	What is measured?	Type of assessment	When to use?
Input	Resources	Planning	Planning of resources
Process	Implementation of activities	Quality assessment	Monitoring of implementation
Output	Effectiveness of implementation	Short-term monitoring	Reporting on immediate progress
Outcome	To which extent did the activities reach their objectives?	Mid-term evaluation	Reporting on intermediate results
Impact	What was achieved by the intervention?	Long-term evaluation	Reporting on real impacts or performance

There is a general consensus that impact indicators are most relevant for the final assessment of success. On the other hand, the combined use of indicators addressing inputs, process, outputs and outcomes helps to capture progress at different time scales (Huovila et al., 2019). Further, the action of discussing indicators, distinguishing short-term and long-term goals, and subsequent monitoring is in itself a societal activity serving (collaborative) policy formulation, planning and implementation (Vermeulen, 2018).

What many practitioners as well as some researchers ignore or underestimate, however, is the growing 'attribution gap' that is inherent in this simplistic, linear input-impact typology. In the case of the tree planting example, the counting of seedlings and monitoring of 'input indicators' is relatively straightforward. However, how can a city know whether long-term changes in air quality have anything

to do with the planting of trees? Perhaps the observed air quality improvements are actually due to other factors (such as national car emission regulations or the closing of a polluting factory in a neighbouring town); it is equally conceivable that air quality has deteriorated yet that the municipal tree planting project actually had positive impacts nonetheless since air quality would even be worse without the additional trees. As shown in Figure 1.5, the complexity of assessment (due to time lags and lower sphere of control) is generally much higher for impact than for input indicators. In terms of language use, it is generally appropriate to talk about the 'measuring' of input indicators if these can easily be counted. This does not apply to impact, however, which can hardly ever be 'measured' in isolation. Instead, impact can only be 'assessed' via a careful analysis of long-term data trends and the study of various contributions including well-intended activities yet also many other confounding factors.

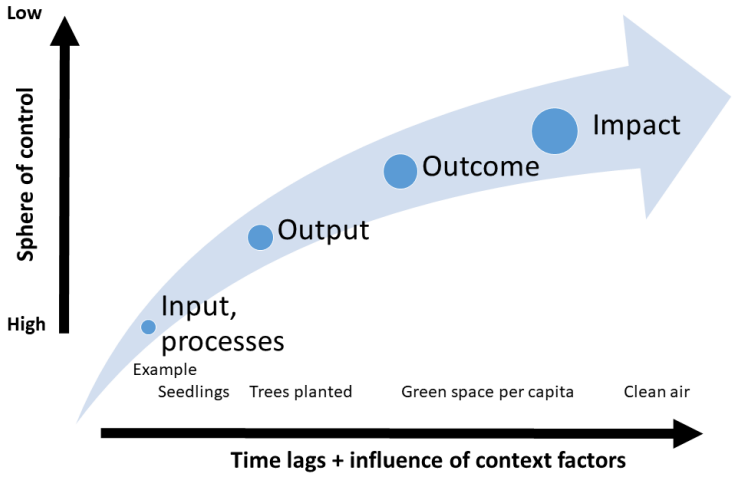


Figure 1.5 *Attribution gap associated with indicator types*

In addition to the input-impact typology – which is highly popular but often misused and underestimated in terms of its complexity – one can of course also distinguish indicators by sustainability-relevant sector (e.g., health, environment, economy)[2]. Furthermore, there is some debate among researchers and practitioners about whether indicators are necessarily quantitative. The SDGs also contain some qualitative indicators, for example about the presence and absence (or degree of implementation) of governmental policies. According to several

[2] In environmental management, the DPSIR model (distinguishing problems and relevant indicators in terms of "drivers, pressures, state, impact, and response") is contested but highly influential (Carnohan et al., 2023).

researchers, policy-based indicators, in general, are underused and under-researched (Dhakal & Imura, 2003; Huovila et al., 2019).

A further point of contention is the role of subjective data sources. The OECD's *Better Life Index*, for example, contains a combination of objective and subjective indicators, for example, homicide rates, and survey data on the question "Do you feel safe walking alone at night in the city or area where you live?" (Zinkernagel et al., 2018). What the OECD also does in its ranking of countries is to integrate both indicators into an aggregate, composite indicator or index called 'safety'. But is such aggregation appropriate? In the logic of all indicators being metrics that simplify complex social phenomena, one can argue that such aggregation is a necessary next step. According to a review article, "there has been a gradual shift in emphasis in the development of the various environmental and sustainable development indicators — from descriptive through performance to composite indicators" (Lehtonen et al., 2016, p. 3). The price, however, is increasing opacity, which some researchers explicitly warn against: "Aggregation is evil when it gives mediatic power to numbers that do not deserve it" (cf. Bell & Morse, 2018b, p. 7).

For the selection of appropriate indicators, another important debate concerns the question of whether indicators should be standardised, and as such defined 'top-down', or developed 'bottom-up', according to local needs and stakeholder input. Researchers observe that in response to the emergence of numerous 'green city' assessments there currently is a "a data standardisation revolution in cities", where "invasive technification" has "tempted many cities to set aside pre-existing participatory and locally tailored data systems to take advantage of Big Data and standardised measures" (Moreno Pires et al., 2017, p. 3). One key advantage of standardisation is that it allows for comparisons and benchmarking; its downside, however, is: i) the risk of being inappropriate for local conditions and ii) the missed opportunity of indicator discussions as a tool for stakeholder engagement and local empowerment (Estevez et al., 2021; Reed et al., 2006). Each city is unique in its context, yet local authorities generally lack resources to develop and maintain tailor-made indicator systems; possible ways to deal with this issue are to apply adjustment weights to standardised indicators or to add bespoke indicators to universal sets if needed (Sharifi, 2019). For the time being, "the literature on standardised frameworks of city indicators is surprisingly scarce" (Huovila et al., 2019, p. 89) and "what constitutes an appropriate balance between the standard aspects of urban sustainability frameworks and the local variation of particular applications remains an open discussion (Bell & Morse, 2018b, p. 7).

It is evident, then, that selecting indicators involves navigating multiple paradoxes, dilemmas, and trade-offs that Lehtonen (2015, p. 92) describes in terms of the following tensions:

- deductive versus inductive approaches;

- use of indicators as inputs for the design and implementation of public policies versus as tools for monitoring and evaluation;
- international comparability versus national/regional/local relevance;
- description versus prescription;
- academic versus practitioner emphasis, i.e. whether an indicator should be defined by its scientific quality or practical usefulness.

For sustainability monitoring (as the ongoing tracking of trends) and reporting (as a one-off or periodic activity), selecting appropriate indicators is a quintessential design step yet arguably overrated. The entire debate is "creator-led" with far too little input from consumers of sustainability indicators (Bell & Morse, 2018b). Too much emphasis goes to the final product, despite growing evidence of the importance of indicator production processes (Lehtonen, 2015). The questions of reporting frequency and authorship – mentioned as key issues of interest in the Introduction (cf. Section 1) – are hardly ever addressed in extant research. However, these issues are important to explore since indicator initiatives are often short-lived. As mentioned above (cf. Section 1.1.2), many urban indicator projects were created in the – probably naive – hope of immediate policy relevance and discontinued when this did not materialise. Many scholars therefore call for other researchers and practitioners to pay much more attention to the question of information usage. According to long-term observers, gradual shifts can be observed in the field of sustainability indicators. Lehtonen (2015) discerned three camps of "true believers, critics and doubters, and pragmatic sceptics". In a similar vein, Holman (2009, p. 370) lamented two main types of research articles that "either engage with a technocratic discourse on indicator development and therefore provide a rather formulaic 'recipe-book' of how to 'do' sustainable development or they emphasise the 'soft', intangible outcomes of indicator programmes without ever coming to terms with how SIs affect policy or alter governance".

To conclude, some critics argue on the grounds of principle that "sustainability indicators are not a convincing alternative to an ethical-legal normativity" (Ekardt, 2024, p. 1); others have begun exploring risks and downsides of indicators (Lyytimäki et al., 2020). At the same time, sustainability indicator systems continue to proliferate. Pragmatists acknowledge that indicators are inherently rigid and partial representations of complex realities (Klopp & Petretta, 2017) but see value in continued experimentation. Experimentation and research is needed concerning productive governance arrangements including multi-stakeholder participation, the choice of indicators including policy-based ones, and actual indicator uses by people in (networks of) divergent cities.

1.2.4 Evaluating performance, accountability and participation

The topics covered so far – sustainability, urban development, and indicators – have brought to the fore several practical explanations and distinctions of relevance for city-oriented sustainability reporting. On this basis, one might draw the lessons that (good) city governance is a complex endeavour and that the mere disclosure of sustainability-relevant data (via indicators) is unlikely to be of immediate and long-term, sustainable, benefit for a city. After reading the preceding sections one might also walk away with the notion that well designed sustainability reports constitute at best a worthwhile initiative, and at worst a harmless waste of effort. After all, reporting potentially contributes to performance, accountability, and transparency, which are generally considered positive for society. However, as a review of public management insights will show, there are many more dilemmas.

The performance of public sector organisations including local governments, has been a major concern for various decades (Bouckaert & van Dooren, 2015). Across Western Europe, North America, and elsewhere, the topic regularly features in public debates and has been at the origin of many institutional reforms since the 1990s. A central component of such reforms (including those known under the label of 'New Public Management') has been the introduction of 'management-by-objectives' or 'results-based management'. This requires the formulation of 'smart' objectives (where the acronym stands for specific, measurable, achievable, relevant, time-bound), the identification of suitable indicators, and, subsequently continuous measurements. This approach is meant to allow organisations achieve three functions (Van Dooren et al., 2015), namely to: i) steer and control (management), (ii) improve (learning), and (iii) justify actions (accountability). This logic has been copied from the private sector and is now embodied in public sector organisations and has permeated the public consciousness in many countries (Pollitt, 2013). International institutions such as development banks (and the UN via frameworks such as the SDGs) also promote it, and the concept resonates with the broader, positivist concept of 'evidence-based policy-making' (Boaz & Nutley, 2015; Cairney & Oliver, 2017; Parkhurst, 2017).

Does it work? According to reviews, the performance management literature is divided about whether performance information and performance measurement actually lead to better quality of service delivery; one key constraint is again – just as in the case of sustainability indicators – the lack of data use by intended users including politicians, professionals and citizens (Van de Walle & Cornelissen, 2014). Well-established, however, is the observation that implementing results-based management is fraught with complexity and brings certain downsides too. At a purely economic level, the costs of data collection and analysis and the proliferation of monitoring and evaluation activities has led to an increase in bureaucratisation (De Bruijn, 2002). More importantly, no matter how well intentioned

a measurement system is, it will inevitably have unintended consequences or 'perverse effects'. In a seminal paper about the "unintended consequences of publishing performance data in the public sector", Smith (1995) identified the following eight phenomena in British institutions[3]:

- Developing a tunnel vision (focusing on measured indicators to the exclusion of other important aspects);
- Suboptimisation (pursuit of local objectives at the expense of organisational objectives);
- Myopia (pursuit of short-term targets at the expense of long-term objectives);
- Measure fixation (an emphasis on measures/indicators rather than the underlying objective);
- Misrepresentation (manipulation of data);
- Misinterpretation (the data is interpreted incorrectly);
- Gaming (manipulation of behaviour);
- Ossification (loss of flexibility and innovation due to rigid indicators).

Measuring (and reporting) public sector performance is thus neither straightforward nor necessarily a benign affair. What about accountability and transparency? The term accountability – which has bookkeeping origins and no exact equivalent in many world languages – nowadays refers to "a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences" (Bovens, 2007, p. 450). Accountability thus extends beyond performance and concerns all types of 'conduct' – both action and inaction[4]. This is highly relevant for (sustainability) reporting because good, substantial reports address and describe not only data but also institutional strategies and achievements.

In this way, however, any well-intended strife for accountability also leads to unintended consequences. Just as in the case of performance measurement, the production of accountability-relevant information (e.g., in reports) has economic costs and may lead to 'accountability overloads' (Halachmi, 2014). The higher the threat of sanctions – an inherent aspect of accountability – the higher the risk

[3] Other researchers have clustered unintended consequences under slightly different labels. Concerning the act of target-setting (as opposed to monitoring and publishing performance data), well-documented phenomena include the ratchet effect (a disincentive to perform above the set target for fear it will be raised), the threshold effect (those already exceeding the target have no incentive to keep over performing), and output distortion (people only do what is measured, even if something else is more effective) (Lewis, 2015).

[4] Since performance measurement is supposed to help organisations render account about their efficient use of resources, some scholars also use the term 'performance accountability' (Van de Walle & Cornelissen, 2014).

that the accountability arrangement stifles flexibility and discourages innovation (De Bruijn, 2002).

Triggered by Anglo-Saxon examples, many countries have witnessed an 'audit explosion' (as coined by Power in 1994) in the previous decades and now have a plethora of mechanisms, including teaching audits, clinical audits, forensic audits, environmental audits (Grossi et al., 2020). Ironically, several reforms were initiated with the explicit objective of increasing (and regaining) public trust in government, yet polls about that matter hardly show long-term improvements. Some scholars go as far to suggest that the transition from a 'representative' to a 'monitory' democracy (Flinders, 2011, p. 595) – based on, for example, constitutional watchdogs, audit processes, ethical guardians, and freedom of information legislation – has actually contributed to the erosion of public support for political processes and institutions.

As explained earlier (cf. Section 1.1.1), city-oriented sustainability reporting is not an exclusive domain of local governments. Instead, it is sometimes initiated and run by civil society organisations or public-private joint ventures. Such modern, 'network-based governance' arrangements bring their own accountability challenges that are underexplored empirically (Mees & Driessen, 2019) but may also help counter the perennial problem of lack of use.

Participation in the design phase is expected to increase feelings of ownership, 'data literacy', and engagement with the emerging results (M. Cohen et al., 2015). There are some documented links between community indicator projects and governmental performance systems (de Lancer Julnes, 2006) and efforts to have citizens participate in performance measurement (Ammons & Madej, 2017; Kroll et al., 2019). Such initiatives can be conceptualised as pertaining to the larger field of 'social accountability' (Joshi & Houtzager, 2012) as well as 'participatory and deliberative processes' (PDPs), which constitute a burgeoning field of experimentation and research in many countries. Participatory approaches are about providing citizens with actual control decisions (e.g., via referenda or participatory budgeting), whereas deliberative methods focus on fostering considered judgments (via community assemblies, citizen's juries, etc.). According to a recent review, "the literature provides mixed and limited results with respect to the ability of PDPs to promote general system-level effects on attitudes and perceptions, such as trust in democracy" (Spada & Paulson, 2023, p. 85). This conclusion of mixed evidence is arguably not surprising and likely to remain valid given the diversity of methods and the heterogeneity of the world's cities (as main field of democratic experimentation), thus precluding grand conclusions.

How do these insights into the into the benefits and downsides of performance measurement, accountability and participation bear on the evaluation of sustainability reporting? In the neighbouring field of citizen initiatives, a systematic

literature review found that many published studies lacked methodological rigour and in particular a clear, theory-informed conceptualisation of effects (Igalla et al., 2019).

Which effects can one – in theory – expect from sustainability reporting? In recent years, many researchers (and seasoned practitioners) have come to realise the complexities of information use. One way to address this issue is to distinguish different types of use. Based on a model developed by Weiss (1979) for research and evaluation findings, Hezri (2004) developed a taxonomy of uses. This model, summarised in Table 1.4, distinguishes 'instrumental use' (rational, positive, tied to decision-making) from conceptual uses (related to learning) as well as three types of less rational uses, including political, symbolic, and tactical.

Table 1.4 *Taxonomy of indicator and information uses*

	Degree of rationality	
Nature of response	High	Low
Positive	Instrumental use • Use for action and decision-making	Political use • Ammunition to support pre-determined use • Discursive use
Ordinary	Conceptual use • Use for enlightenment and learning	Symbolic use • Ritualistic assurance
Negative	Non-use	Tactical use • Delaying or substituting action • Deflecting criticism

The boundaries between types of use and influence are, of course, not clear-cut; what one may genuinely regard as 'instrumental' may another regard as 'political'. However, the model has become very influential in research on sustainability indicators and monitoring initiatives (Bell & Morse, 2018b, p. 7). Some scholars have proposed expanding the notion of 'use' by also including the notion of 'influence' (Lehtonen et al., 2016), and clustering the main (potential) benefits of sustainability monitoring in three domains: i) instrumental use and influence (decision-making), ii) conceptual use and influence (learning), and iii) political-symbolic use and influence (discourse, agenda-setting). Based on performance management and accountability research, one can expect main drawbacks of reporting in the domains of rigidity, loss of flexibility and innovation, and bureaucratisation in general.

To date, very little sustainability research has explored the actual use of indicators in political decision-making. There are some conceptualisations about sustainability indicators in relation to: i) the 'policy cycle' model (Lehtonen, 2015; Seaford, 2013) and ii) municipal management (Krantz & Gustafsson, 2021; Park & Krause, 2021; Zeemering, 2018). Hardly studied empirically, however, are the benefits and downsides including 'perverse' effects of sustainability indicators in political decision-making and performance management of local governments.

In this context, extant research has also made insufficient use of more fine-grained insights and models emerging from the study of performance management in different policy sectors. More than 30 years ago, Wilson (1991) proposed a typology of public sector organisations according to the 'observability' of their outputs and outcomes. In this model – visualised in Table 1.5, adapted from van Dooren et al. (2015) – tax authorities are examples of organisations with observable outputs (e.g., the amount of tax filings) and observable outcomes (e.g., overall revenues). The work of mental health counsellors can also be related to certain outputs (e.g., the number of patients attended), but their desired outcomes (e.g., wellbeing) are difficult to observe and attribute. In the case of general practitioners offering a range of services, concrete outputs are hard to pin down yet certain outcomes (e.g., disease level of patients) are measurable. For diplomats and researchers, neither outputs nor outcomes are generally observable. Another approach deserving replication is the screening of indicators in terms of 'understandability' (by lay people) and the susceptibility to gaming and cheating' (by professionals), as applied by Pollitt (2011) in one study on the PISA scores and the WHO's life expectancy rates. These models are – as all models – a simplification of reality yet offer valuable insights for different organisations and indicators; the better the observability and understandability of outputs and outcomes, the less problematic the application of performance management (and sustainability reporting) systems.

Table 1.5 *Typology of organisations according to Wilson*

		Outcomes observable?	
		Yes	No
Outputs observable?	Yes	Production organisations (e.g., mail services, tax collection, sanitation)	Procedural organisations (e.g., military (peacetime), mental health counselling, youth penitentiary)
	No	Craft organisations (e.g., military (wartime), doctors, forest rangers)	Coping organisations (e.g., diplomacy, intelligence, research)

To sum up, sustainability reporting as a tool of (public) management and accountability is likely to have numerous positive, intended effects as well as unintended and potentially negative ones. Main potential benefits can be expected in the domains of better decision-making, learning, and agenda-setting, and costs in the realm of rigidity and 'accountability overloads'. In evaluation research, many variables need to be taken into account such as the observability and 'lay' understandability of indicators (per sector and policy area) as well as people's 'data literacy' and degree of participation.

1.2.5 Private sector accounting and reporting

Some casual observers of sustainability reporting make the offhand observation that the practice is well developed in the private sector and underdeveloped in the public sector, implicitly suggesting that the latter should simply emulate the former. These comparisons are overly simplistic. Private sector, for-profit organisations have a different *raison d'être*, and their overarching goal of 'maximising shareholder value' can always be related to a monetary 'bottom line'. By contrast, the public sector has far more diffuse and complex roles, values, stakeholders, responsibilities and criteria for success or performance (Kaur & Lodhia, 2019). As already mentioned (cf. Section 1.1.1), public sector organisations also tend to emit more information in more diverse ways (via for example, periodic reports, projects reports, websites, or dashboards), which raises the question of what exactly counts as a sustainability report and what does not. By comparison, for business companies one can operationalise and determine more easily whether they do or do not engage in the production of (annual) sustainability reports or 'integrated reports' that include sustainability dimensions.

Despite these systemic differences, reviewing private sector developments in terms of accounting practices and standards remains relevant for cities. The private sector simply has a longer reporting history, a much larger number ('n') of organisations studied, and a larger body of research producing insights of relevance for sustainability reporting in general. One recent review examined 1,283 academic articles published in 54 journals between 2014 and 2020 and found that sustainability disclosure is the most frequently researched topic, studied mainly via archival (i.e., desk research) methods, with a substantial proportion of publications analysing listed firms in a national, predominantly European setting (Hsiao et al., 2022).

Other frequent research topics include questions of motivations, i.e., why organisations engage voluntarily in sustainability reporting. In this context, a common critique is that firms' actions are driven by impression and reputation management, aiming to maintain social legitimacy rather, than a genuine commitment to sustainability. For others, this is not a contradiction as practices can be both symbolic and substantive (Sheehy & Farneti, 2021).

A further explanation for the spread of reporting practices is 'institutional isomorphism', which includes normative, coercive and mimetic forms (Pollitt, 2003). In terms of norms and coercion, external pressure plays a role since several investors, banks, and rating agencies have revised criteria for business performance (Cardoni & Kiseleva, 2023). In Europe, the Corporate Sustainability Reporting Directive (CSRD) came into force in 2023 (replacing the Non-Financial Reporting Directive of 2014) and requires almost 50,000 business companies to disclose information on their business model and strategy, policies, risks, targets, and due diligence in relation to "environmental, social, and governance" (ESG) issues (Cardoni & Kiseleva, 2023).

Companies use a variety of standards including those of the Global Reporting Initiative (GRI), the Integrated Reporting Framework (<IR>), and the Sustainability Accounting Standards Board (SASB). However, the evolution of the GRI serves as a cautionary tale regarding the role of civil society. Levy, Brown and de Jong (2010) analysed its political trajectory and found that while the initiative was originally intended to empower civil society, the institutionalisation process was largely shaped by corporate power dynamics, ultimately prioritizing financial logic over genuine social accountability.

Numerous firms also relate to the SDGs (Bebbington & Unerman, 2018). For some scholars, framework diversity is a strength rather than a weakness (Bose, 2020). There also is some overlap in terminology inasmuch as the buzzwords 'corporate social responsibility' (CSR), 'corporate sustainability' and 'ESG reporting' are often treated as synonyms (Sheehy & Farneti, 2021). However, treating these terms as interchangeable is problematic, as neither companies nor cities can be 'sustainable' in isolation.

The widespread uptake of sustainability reporting (or, more precisely, of 'ESG disclosure') by commercial companies has had the practical effects that numerous staffers (and their accountants) have gained practical experience with frameworks, indicators, and the application of abstract formats to real-world organisations. In this context, formulating one's organisation's 'scope' is an important step and therefore a key object of discussion. This refers to defining an organisation's reach and influence, and connects directly to the challenge of delineating outcomes in performance-based management (cf. Figure 1.5 on the 'attribution gap' of the results chain logic).

For the business sector, Sheehy and Farneti (2021) elaborated a visual model (cf. Figure 1.6) showing that firms have some leeway in defining their scope as purely organisational ('weak corporate sustainability') or looking into the spheres of national and international influence ('strong corporate sustainability'). This line of thinking may also be helpful for sustainability reporting at the city level, where discussions of scope play out in a similar fashion. Another generic technological

development that is currently studied in the corporate sector and likely to affect sustainability reporting everywhere concerns the benefits and pitfalls of AI text generation (de Villiers et al., 2024).

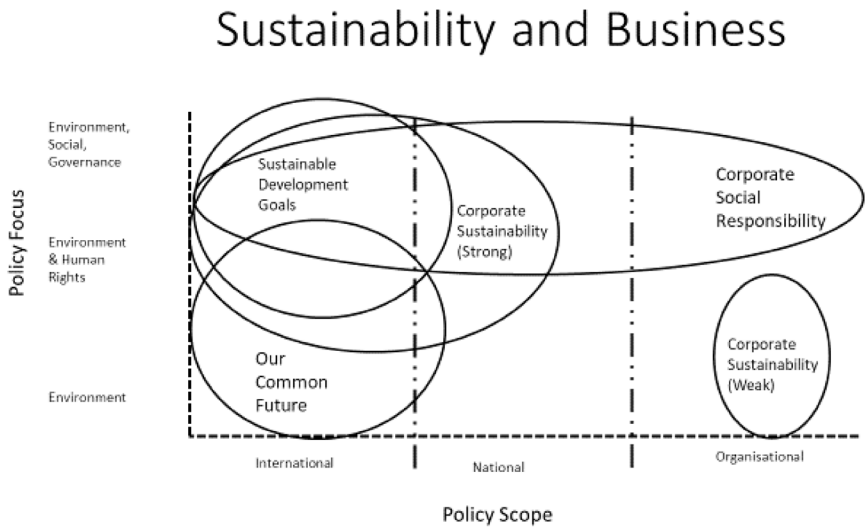


Figure 1.6 *Scope and policy focus options for firms*

Adapted from "Corporate social responsibility, sustainability, sustainable development and corporate sustainability: What is the difference, and does it matter?" by Sheehy & Farneti (2021)

A crucial question concerns, of course, the outcomes and consequences of reporting. Research on private sector reporting typically focuses on three effect clusters: (i) financial performance, (ii) internal operations and management practices, and (iii) external perceptions of accountability/legitimacy. According to a recent review by Hsiao et al. (2022), many studies have documented a positive association between sustainability accounting practices and a firm's monetary value. Concerning internal affairs, research findings typically support the argument that the implementation of environmental management systems improves environmental performance and environmental capabilities (Velte, 2022); also in firms, however, some studies found that sustainability information is underused and does not result in transformative changes in operations. On the question of the extent to which sustainability reporting benefits reputation and legitimacy, the evidence is mixed (Velte, 2022). As already observed 15 years ago, the phenomenon of 'reporting fatigue' can set in when disclosure becomes institutionalised, causing organisations to view the process as a burden rather than a strategic opportunity for change (Brown et al., 2009).

Should policymakers and regulators push for mandatory reporting? A major empirical study addressing this question recently produced a warning. The EU's directive 2014/95/EU on non-financial reporting was introduced with the explicit objective of improving corporate social and environmental performance (rather than abstract legitimacy) and was studied as a natural experiment. De Villiers et al. (2025) assessed how 358 European companies complying with the EU directive performed over the long term and compared them on the basis of various indicators to 470 US American counterparts lacking such reporting obligations; the comparison showed no significant increase in social or environmental performance in the European sample.

To conclude, evaluating the practice and in particular the effects of sustainability reporting in the private sector is different from city-oriented, public sector sustainability reports that take on much more diffuse forms and may serve more different purposes. Nonetheless, there are opportunities for cross-learning and cross-fertilisation, as certain aspects (e.g. the challenge of information use) play out in both domains.

1.2.6 Comparative politics and adaptive policymaking

Most models discussed thus far focus on entities such as local governments, companies, and civil society organisations, implicitly treating these institutions as stable, legitimate units of observation and analysis. Much of the literature concentrates on what these organisations produce—or fail to produce—in terms of sustainability reports, often with the unspoken assumption that current findings will remain valid in the near future. However, two fundamental assumptions underpinning this work require closer scrutiny: that institutional categories are universally applicable, and that findings hold consistently over time. In social science research, it is necessary to construct models using abstract variables and concepts. At the same time, generalisations must be approached with caution so as not to obscure other critical factors. In the case of city-oriented sustainability reporting, two such under-addressed factors are contextual diversity across cities and temporal change.

The challenge of comparative diversity requires recognising that the world's countries—and their cities—are too heterogeneous for universally applicable theories or recommendations to be readily extracted. Acknowledging this, and noting the strong publication bias favouring studies in Western, Educated, Industrialised, Rich and Democratic (WEIRD) countries (Kahalon et al., 2022), many scholars call for expanded research in underrepresented regions. Meier et al. (2015, p. 20), for instance, contend that, "to advance our understanding of the different assessments of transparency, the framework needs to be applied differently in different countries. Previous work suggests that transparency has a different impact in different institutional and cultural settings". Similarly, Grimmelijkhuijsen and Kasymova (2015), in their comparison of the Netherlands and

the United States, emphasise the importance of regime type—arguing that in majoritarian democracies, transparency serves to empower excluded interests, whereas this is less pressing in consensual systems.

Cultural context also matters. Hood (2012), for example, suggests that performance management through target-setting works best in 'hierarchist' cultures but poorly in those dominated by 'fatalist' worldviews. At the macro level, Pollitt and Bouckaert's (2011) typology of politico-administrative systems—grouping countries into categories such as Anglo-American, Napoleonic, or Germanic—offers useful, though still partial, frames of comparison. These models are most applicable to high-income nations, but certain underlying constructs (such as Weberian notions of bureaucratic autonomy or degrees of managerial discretion) can inform broader global research.

Urban governance in lower-income contexts presents a distinct set of challenges often overlooked in wealthier settings. These include unreliable data systems, informal institutions (e.g., unregulated transport systems), and a scarcity of resources to support systematic reporting (Jerven, 2013; Klopp & Petretta, 2017). In the literature on smart cities, one can also find pleas for "greater attention to reform in lower-income and corruption-riddled contexts" (Grossi et al., 2020, p. 634). Moreover, transparency and civic space are themselves politically contested. With democratic backsliding in various regions—particularly in the Americas—the role of the state and public access to information are increasingly fraught (Mendonça & Grandé, 2023).

From a temporal perspective, the presence or absence of sustainability reporting (understood here as a periodic and public activity; see operational criteria in Table 1.1) is relatively easy to assess. Consequently, many indicator-based studies have tracked the disclosure behaviours of local governments (Guerrero-Gómez et al., 2021; Navarro Galera et al., 2014; Navarro-Galera et al., 2017). While these studies offer a descriptive foundation for international comparisons, they do little to illuminate how reporting evolves or what effects it produces over time.

Longitudinal studies are needed to address such dynamics. Some of this work is emerging through institutionalist lenses, such as research on the "institutionalisation of evaluation" (Stockmann et al., 2022). However, there is a risk of equating institutional expansion with improved quality. The same applies to the proliferation of 'maturity models' that seek to assess local governments' progress in areas like data governance (Kroll, 2015), open government (Ingrams, 2020; Pirannejad & Ingrams, 2022), and managerial practices (Rosvall et al., 2023). While such conceptual frameworks can be analytically useful, they are often difficult to apply rigorously to complex, normative concepts like governance (Heywood, 2015). Moreover, in international settings, the application of maturity scales or governance ratings is hotly contested—as is the broader pursuit of

identifying 'best practices' (Andrews, 2010; De Vries, 2010; Guinn & Straussman, 2018). Indeed, the concept of 'good governance' that dominated policy debates in the 1990s has largely fallen out of academic favour.

In light of these limitations, adaptation may offer a more fruitful conceptual approach to understanding sustainability reporting. Defined as "the process of changing to suit different conditions" ('Adaptation', 2024), the notion shifts attention from static comparisons or normatively loaded benchmarks to context-sensitive change and responsiveness. Adaptation implies attentiveness to complexity, emergent phenomena, and local specificity. Various adaptive approaches have been proposed in recent literature, including adaptive policymaking (Walker et al., 2001), adaptive governance (Janssen & van der Voort, 2016), and "problem-driven iterative adaptation" (Andrews et al., 2017)—a framework that has gained popularity in development studies. Common features of these approaches include experimental policy implementation, rapid feedback loops, anticipation of uncertainty, and deep sensitivity to contextual variation.

To summarise, future research on sustainability reporting should prioritise comparative, longitudinal, and adaptive approaches. Addressing geographical and temporal blind spots is critical for generating more robust, transferable knowledge. Rather than pursuing elusive best practices, scholars and policymakers should focus on how reporting practices adapt to local conditions and change over time.

1.3 Knowledge gaps, research questions, and methods

The preceding sections have introduced 'city-oriented sustainability reporting' as an empirical, real-world phenomenon (Section 1.1.) and, as a topic of research and theoretical interest (Section 1.2). These introductory sections provided an overview of the background and what is already known about the subject matter and where other scholars see research gaps. On this basis, the current section takes stock and synthesises these gaps in order to formulate appropriate research questions. Importantly, identifying research questions is not a matter of passive observation or ideas simply falling into place. Rather, it requires an active process of matching real-world opportunities with identified knowledge gaps.

In response to the need to consciously carve out the research design, the description of knowledge gaps (Section 1.3.1) is followed by a description of the rationale and purpose of this dissertation (Section 1.3.2). This leads to the formulation of research questions (Section 1.3.3.). To answer these questions, each of the four empirical studies conducted with a tailor-made selection of study material and methods for data collection and analysis (cf. Chapters 3-6 containing a 'methods section' for each study). While each empirical study is unique, there is also a unifying methodological 'red thread' running through the dissertation. This is

explained in the final sections of this introductory chapter in terms of three commonalities: i) All studies are longitudinal, comparative case studies (Section 1.3.4), they ii) all employ research frameworks with common elements (Section 1.3.5), and iii) all apply a 'mixed methods' research design (Section 1.3.6).

1.3.1 Gaps in knowledge and extant research

The literature review in Section 1.2 highlighted a number of research gaps and open questions as raised by prior scholarship. These gaps can be broadly categorised into two areas: theoretical and empirical.

- **Theoretical gaps**

The literature review showed that 'urban sustainability' has a global appeal as policy aim yet is difficult to pin down conceptually (Section 1.2.1) and a complex matter in terms of policy implementation. Urban governance requires alignment in terms of vertical coordination (from citizens to the international level) and horizontal cooperation (between local governments and many non-governmental stakeholders). Sustainability reporting is a popular tool of 'informational governance' (Section 1.2.2) and likely to have positive effects as intended by its proponents as well as negative or 'perverse' ones (Section 1.2.4). What the current literature does not contain are answers to the question which design choices in terms of reporting content and internal governance arrangement are associated with which type of emerging use, benefits and drawbacks in which type of geographical and temporal context. Put differently, the four key knowledge gaps about (city-oriented) sustainability reporting concern the production processes (including content and authorship), their contextual embeddedness in time and place, and their varied effects—both positive and negative.

According to a recent review (Sharifi et al., 2024, p. 146), "at the city level, there is a need for extensive research, especially on the best planning models that could be adopted to facilitate the achievement of the sustainability goals". As already mentioned, questions of governance are also under-researched. Most existing studies are silent on the question of who is involved in city indicator and reporting initiatives. One study observed that European initiatives tend to be run in cooperation with governments and North American ones 'at arm's length' (Holden, 2013). In this light, attempts to compile 'best-practice' recommendations such as the prescription that indicator initiatives should "include a balanced mix of government, business and community representation" (Davern et al., 2016) appear rather uncritical. According to other literature reviews identifying future research directions, "the partnerships between the public sector and other entities (such as business groups or even non-governmental organisations) in relation to sustainability issues and the resulting impact on sustainability accounting, reporting

and accountability is another worthwhile research endeavour" (Kaur & Lodhia, 2019, p. 500).

One major theoretical gap concerns the question of standardisation. Along with the mainstreaming of 'urban sustainability' as a global policy aim, there are many calls for standardisation and upscaling, for example by the United Nations and ISO (ISO, 2014). The main (and generally plausible) argument in favour of standardisation is that shared frameworks and methods stimulate performance and facilitate the transfer and diffusion of successful practices (Castellaci et al., 2005). Comparisons and benchmarking are important practices in this regard. However, what exactly should and can be standardised is contested – most organisations and scholars focus on the reporting of sustainability outcomes and assessments (i.e., sustainability indicators), suggesting that policy processes should be left locally determined. This approach fits the paradigm of outcome-oriented performance management widely disseminated by New Public Management (NPM). At operational level, the link between sustainability monitoring and accountability constitutes an important, underexplored design question (cf. Section 1.2.4).

Others, however, pursue the opposite line, arguing that universal indicators and targets are unrealistic in the face of highly divergent historical and socioeconomic realities (Joss, 2012). According to this line of reasoning, focusing on sustainability-oriented processes is more productive. Adherents contend that, "what is most needed to accelerate and scale up innovation – particularly in the context of rapid urbanisation in developing global regions – is a process focus, prompting the need for generic, replicable protocols and tools for supporting the design, implementation and assessment of urban sustainability initiatives" (Joss et al., 2015, p. 9). In this approach, an operational issue worth researching are the advantages and disadvantages of making the use of tools mandatory or voluntary for cities. These contradictory assertions underscore a core theoretical gap that merits further investigation. The tension between comparability and contextual relevance lies at the heart of city-oriented sustainability reporting. Rather than treating standardisation and tailoring as mutually exclusive, this dissertation explores how hybrid approaches—where local adaptations are built upon broadly comparable frameworks—can provide a workable balance.

- **Empirical gaps**

Two main empirical shortcomings stand out in the literature on city-oriented sustainability reporting. The first concerns the geographical distribution of research. Urban governance studies suffer from a significant publication bias; the vast majority of studies concern developments in a select, over-represented number of world regions, notably Europe, North America, and Australia (Ferreira da Cruz et al., 2019). In the previous decade, some indicator work has been done on Asian cities (Krank et al., 2013; Krank & Wallbaum, 2011), but developments and

practices in African and Latin American cities have hardly been studied. This poses major problems for the validity of models and policy prescriptions inasmuch as governance is context-dependent. For practitioners designing sustainability interventions, Meuleman and Niestroy (2015, p. 12314) recommend that "on each relevant administrative or geographical level, governance design should begin with stock-taking of the governance environment, taking into account the specificities of each nation/region/city, i.e., analysing the existing governance arrangements including what has worked historically and where are the gaps and obstacles". A systematic literature review of democratic innovations (Igalla et al., 2019, p. 1188) also concludes that, "the field of citizen initiatives could use more comparative research (e.g., in different countries, continents, and among sectors)".

The second empirical gap relates to the evolving use of sustainability reports. Much extant research is limited to cross-sectional snapshots, merely assessing disclosure patterns based on the analysis of websites and the manifest content (i.e., explicit and surface-level information) of publicly accessible reports. While this method is efficient and suitable for certain questions, it fails to capture the evolving use of information over time. In their influential book on sustainability indicators, Bell and Morse (2018b, p. 12) assert that "interest groups will always make selective and distorted use of indicators… Thus, we would like to see a greater emphasis in research on the space between production and use". This observation aligns with the following critique from a public management perspective (Zeemering, 2018, p. 6):

> We remain uncertain about the extent to which sustainability performance indicators contribute to social learning processes and organisational change. Often, research assesses the adoption of performance measures, and not the management process of performance measure use. To state this critique clearly, sustainability strategy and the management of organisational processes for the pursuit of sustainability within the organisation are often neglected.

Presumably for economic reasons, most studies also examine sustainability reporting initiatives by taking a cross-sectional snapshot of disclosure (or possibly use and governance arrangements) at one moment in time. Considering evidence that indicator initiatives often change and are discontinued, a longer time frame exploring the evolution and adaptation of practices is likely to produce valuable insights. A study on city planning methods (Giles-Corti et al., 2016, p. 9) concluded that, "stronger longitudinal evidence is needed across the board, particularly from natural experiment studies of policy interventions that would allow policy impacts to be monitored to learn what works in different contexts, particularly in LMICs (Low and Middle Income Countries) experiencing rapid changes in patterns of urbanisation". As municipal sustainability reporting is largely

voluntary, it can also be perceived as an innovation to which the following statement applies: "The question whether public sector innovations last, and what determines their chances of survival, remains a gap in the public management literature" (van Acker & Bouckaert, 2018, p. 1).

In summary, the literature on city-oriented sustainability reporting remains both theoretically underdeveloped and empirically narrow. These shortcomings are interlinked: limited geographical coverage weakens the empirical grounding of existing theories, while those theories often fail to capture context-specific dynamics in less-studied political settings. How can we remedy this? In his review about "challenges and opportunities" for research on sustainability indicators review, Ramos (2019, p. 16) identifies ranks the following three approaches as most promising: (i) "A richer selection of case studies", (ii) attention to the "specific cultural context", and (iii) analysis of the most "adequate level of standardisation". Bell and Morse (2018b, p. 5) add to this the consideration that, "what we need is a meta-analysis of sustainability indicator use experiences, but, to do this, we need the case studies to be peer reviewed and placed in the public domain". In line with these suggestions, this dissertation sets out to carry out comparative case studies on i) underexplored sustainability reporting practices including issues of internal governance, indicators and accountability and ii) multiple cities from underrepresented world regions.

1.3.2 Research purpose and rationale

This dissertation pursues a threefold purpose: (i) to contribute to urban sustainability theory, with a particular focus on the use and usefulness of sustainability reporting; (ii) to generate empirical knowledge about reporting practices in under-researched world regions; and (iii) to offer practical, actionable insights relevant to international policymakers and urban practitioners.

The central intellectual puzzle that guides this study concerns the causal factors that help explain which types of sustainability reporting systems—defined in terms of both content and internal governance—are most effective in specific settings. These 'settings' refer not only to broader legal and political systems but also to local discourses on sustainability and the degree of stakeholder involvement. Insights into what works, under what conditions, and why can inform more deliberate design choices and clarify the trade-offs between different reporting approaches.

This knowledge is particularly relevant for civil servants and activists engaged in reporting initiatives, as well as for policymakers who may consider supporting—or mandating—specific reporting mechanisms. In this context, identifying design strategies that are (i) tailored to diverging purposes, (ii) capable of producing desired outcomes while minimising unintended consequences, and (iii) sustainable over time, is of both scholarly and practical significance.

However, two important cautions must be acknowledged. First, political systems and governance arrangements are dynamic. The effectiveness of any reporting approach must be evaluated with awareness of such ongoing evolution; what works in one year may not be as effective in the next. Therefore, regular reassessment of design choices is essential for practitioners. Second, in terms of theoretical ambition, this dissertation does not aim to construct a universally applicable model. As noted by Merton in his seminal work on middle-range theories, social science research often oscillates between narrow case studies and overly general "grand theories" (Yang & Hsieh, 2007). This study aims to contribute instead to middle-range theorising: generating insights into sustainability reporting as a social mechanism that are grounded in empirical research, yet conceptually transferable across similar contexts.

This approach requires strategic choices about research scope and design. As outlined in previous sections, the field is marked by both valuable theoretical contributions (Section 1.2) and persistent knowledge gaps (Section 1.3.1). In line with the funnel metaphor of research design (see Figure 1.7), this dissertation narrows its focus through two key decisions.

The first is a comparative case study design focusing on underexplored sustainability reporting practices across divergent socio-political systems. In response to calls for more comparative research, the study deliberately selects cases from two world regions—Europe and Latin America—that share certain democratic ideals but differ in institutional capacity and developmental trajectories. European cities are typically embedded in high-capacity bureaucratic systems with established sustainability policies, while Latin American cities, despite lower human development indicators, display high levels of urbanisation and a notable "experimentalist vocation" in democratic innovation (Pogrebinschi & Ross, 2019). Politically, Europe is largely characterised by consensual democracies, whereas Latin America tends toward majoritarian regimes (Lijphart, 2012).

The second strategic decision is to focus on the production, costs, and benefits of sustainability reporting, adopting a longitudinal perspective. Since the development and use of sustainability indicators is an ongoing and evolving process, a snapshot approach would be insufficient. A longer-term lens is particularly appropriate for studying stakeholder involvement—an important but under-researched design dimension. Furthermore, this dissertation goes beyond documenting disclosure patterns by investigating reporting effects in three domains: decision-making, learning, and public discourse (cf. Section 1.3.5). While this broad ambition introduces methodological complexity (cf. Section 1.3.6), it is justified by both gaps identified in the literature and calls for more nuanced studies.

In particular, the dissertation places emphasis on two critical themes:(i) city-wide trend monitoring through sustainability indicators, and (ii) the assessment of

local government performance via accountability frameworks and policy indicators. Both areas are crucial to understanding the practical value of reporting and offer promising ground for empirical and theoretical advancement.

1.3.3 Research questions

Sustainability reporting in urban contexts can be understood as a transparency and accountability tool. Accountability, by its nature, is a dialogic process—one that involves posing and responding to questions. In line with this perspective, and following long-standing traditions of inquiry such as the Socratic dialogue, this dissertation is structured around a set of explicit research questions that guide its empirical and theoretical contributions. The central research question is as follows:

> **How can sustainability reporting initiatives benefit decision-making, learning, and public awareness in the divergent contexts of European and Latin American cities?**

To operationalise this inquiry, and based on the rationale for a comparative and longitudinal case study design (cf. Section 1.3.2), four sub-questions have been formulated:

1. **How have sustainability reporting practices evolved in pioneering European city governments, and what are their effects?**

2. **How have sustainability reporting practices by community indicator initiatives evolved in selected Latin American cities, and what are their effects?**

3. **What are the strengths and limitations of reporting initiatives addressing the sustainability policies of selected European local governments via voluntary benchmarking?**

4. **What are the strengths and limitations of reporting initiatives addressing the sustainability policies of selected Latin American local governments via accountability obligations?**

These questions structure the empirical investigations across the dissertation and enable comparative analysis of how sustainability reporting functions as a tool for governance in diverse socio-political environments.

1.3.4 Commonality one: Longitudinal case studies

As outlined in Section 1.3, the four empirical studies in this dissertation explore different aspects of urban sustainability reporting, yet they share a common methodological feature: the use of longitudinal case studies.

To research how some 'transparency and accountability tool' plays out in different contexts, one could, in principle, conduct an experiment: Find cities willing to

participate, agree on a common approach, implement it, and evaluate it. For developing countries, such kind of 'action research' is not entirely uncommon. The World Bank and the International Monetary Fund often condition support to measures of financial openness and push for participatory anticorruption initiatives (Ingrams, 2020) and social accountability mechanisms (Fox, 2015; World Bank, 2017). To trial and evaluate the effectiveness of social policies, randomised controlled trials (RCTs) have also been implemented (Peters et al., 2018). Such approaches certainly have their merits but also limitations; after all, participation in experiments affects one's incentives and motivations, and standardised policies disregard local contexts.

Instead, this research adopts naturalistic case studies, which allow for in-depth understanding of specific practices over time. In research, a case study usually concerns a single phenomenon while one still wishes to generalise findings to a larger set of similar cases (Yin, 2018). The advantage of this approach is gaining in-depth knowledge and doing justice to the peculiarities of the particular case, which are described with accuracy; the downside are external validity concerns when one tries to extrapolate findings to other cases. Hypothesis testing is also not appropriate in general as case studies mainly serve to generate hypotheses. In large-N studies covering numerous cases, the pendulum swings the other way as generalisations have a larger evidence base whereas the peculiarities of individual cases may be underestimated and individual data points may be less reliable.

To balance these trade-offs, each of the four studies in this dissertation is designed as a cross-case comparative analysis. For two thematic areas—trend reporting and local government performance—cases were selected from relevant European and Latin American practices, using operational criteria outlined in Section 1.1.1 to ensure minimal comparability.

Urban sustainability reporting remains a relatively recent phenomenon, and this study focuses on longitudinal effects rather than snapshot evaluations. As a result, early adopters are likely overrepresented. According to Bell and Morse (2018b, p. 5), studying them via "case studies can provide early examples of experiences which may become general trends, weird results which provoke curiosity, even contradictions to the established opus of 'truth'". Studying frontrunners —sometimes termed "positive deviants" (Herington and Van de Fliert, (2018) is a well-established research practice yet needs to be supplemented by other studies for upscaling (Fuhr et al., 2018).

1.3.5 Commonality two: A basic conceptual framework

To adequately assess each case study, the reporting practice must first be analysed on its own terms. As Joss and Rydin (2018, p. 364) argue, "the application of urban sustainability indicators and standards necessarily has to be analysed

within particular local settings; at the same time, it requires attention to wider governance dynamics." Accordingly, each empirical study in this dissertation is grounded in a custom conceptual framework and methodological approach. Despite contextual differences, all four studies investigate three core dimensions: (i) the context, (ii) the design of the reporting initiative, and (iii) its effects.

The context is described in terms of the city's institutional background, concurrent reporting practices, and changes over time[5]. One important design choice concerns the internal governance arrangement—specifically, whether the reporting initiative is led by a local government or by a civil society organisation. A second key design aspect relates to content, including indicator selection and the modes of production and dissemination (i.e., how and through which channels the report is made public). The third dimension—effects—is particularly challenging to evaluate. Presumed benefits, such as accountability, are often intangible, and observed changes can rarely be attributed to a single factor. Causal claims are thus difficult to sustain.

Although the research is guided by this explicit theoretical framework, it is important to note that several key insights—particularly regarding the 'ecological niche' of reporting—were refined inductively through longitudinal comparison across cases and over time, rather than being fully articulated *ex ante*.

The term *impact* is now widely used—sometimes indiscriminately—by activists and researchers alike. North American indicator projects, for example, often attempt to assess their 'community impacts' (Sirgy, 2022; Wood, 2016). As noted earlier (cf. Section 1.2.3), *impact* was traditionally reserved for large-scale, systemic change. The term *results* is somewhat more neutral, but still implies a degree of success or achievement that may not always be warranted. Social innovations often generate unexpected or negative outcomes, frequently referred to as *perverse effects*. For these reasons, this dissertation adopts the more generic term *effects*, which allows for both intended and unintended consequences[6].

---

[5] For the study of governance mechanisms, other research frameworks such as Wittmayer et al. "analytical lenses" (Wittmayer et al., 2015) also assess the "vision" and "aims" of each approach. With regard to sustainability reporting mechanisms, questions regarding aims and expectations are certainly worth asking to local stakeholders but are not necessarily a variable of analysis. After all, as reported in the literature, sustainability reporting is sometimes started as a "fashion" (Biondi & Bracci, 2018) or in response to mimicry and external pressure, without any explicit objectives.

[6] The term *effect* also warrants reflection, as the related concept of *effectiveness* typically refers to the achievement of stated objectives—potentially overlooking unintended consequences. One way to address this limitation is by introducing more generic constructs such as *performance*, which is often conceptualised as a multi-dimensional indicator encompassing effectiveness, efficiency, innovation, democratic quality, and other factors (Edelenbos et al., 2021). An alternative approach is to frame research questions explicitly in terms of both positive and negative effects. This dissertation adopts the latter strategy:

What types of effects can be expected? Based on the literature review, three main domains are considered: instrumental, conceptual, and political-symbolic use and influence of sustainability reporting—as well as non-use (Lyytimäki et al., 2013, 2020). In terms of users, it is useful to distinguish among (i) individual-level effects, such as changes in knowledge or skills; (ii) group-level effects, such as improved cooperation between organisations; and (iii) societal-level effects, such as shifts in public discourse and awareness (cf. Spada & Paulson, 2023). Figure 1.7 summarises the main elements of the evaluation framework, which includes two contextual factors (place and time), two design dimensions (internal governance and content/dissemination), and three types of use and influence. This framework underpins all four empirical studies in the dissertation.

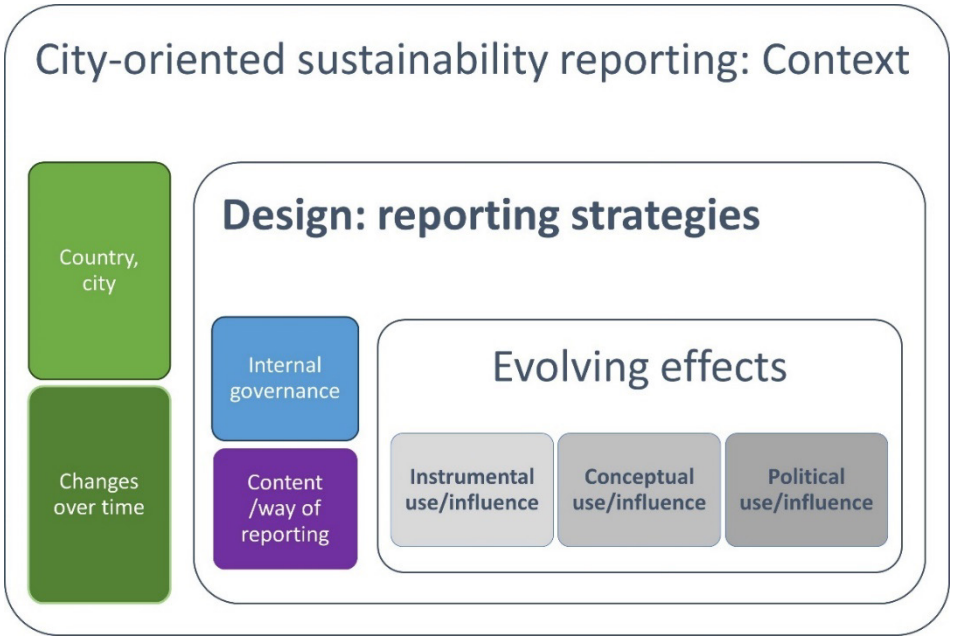


Figure 1.7 *Evaluation framework linking context, reporting strategies and effects*

1.3.6 Commonality three: Mixed methods

To answer the research questions, this dissertation combines conceptual constructs with multiple methods of data collection and analysis. As described above, the empirical studies assess three sets of variables: (i) contextual factors, (ii)

two research questions refer to "strengths and weaknesses," thereby guiding the inquiry toward identifying both desirable outcomes and potential drawbacks or inefficiencies, including economic ones.

design characteristics of the reporting initiative, and (iii) longitudinal effects in terms of various forms of use and influence.

While value-laden concepts such as culture, maturity and governance are difficult to 'operationalise' for comparison purposes, this is even more challenging for the intangible issue of information use and thus the effects of sustainability reporting. How can one assess whether municipal employees or citizens, for example, learn from a report? Researching how a report is produced and used – or perhaps ignored or misused (Lyytimäki et al., 2013) – by a person, group of people, organisation, or society – is infinitely complex and never traceable with accuracy. In the absence of validated frameworks and grand theories, more descriptive research based on qualitative methods may be necessary to advance our understanding of basic governance mechanisms (Sáez, 2012).

As explained earlier, in the context of Wilson's typology of organisations (cf. Table 1.4), some outcomes or effects are observable while others are not. This also applies here. When a sustainability report contains a written justification concerning municipal policies in relation to certain indicators, for example, this may constitute formal evidence of 'instrumental' use. By contrast, questions of motivation and learning can only be assessed indirectly by asking or observing people. When direct evidence is not available, using circumstantial or anecdotal evidence or so-called 'proxy indicators' is the only way forward, yet one ought to acknowledge their limitations (Browne, 2022).

One can thus expect different types of evidence for each of the variables: Information about the context at the country or city level may, to some extent, be derived from other literature and "governance indicators" that one can refer to; the context at institutional level (i.e. a local government or civil society organisation) probably requires querying local stakeholders. Since the reporting practice takes by definition place in the public domain, it can be partly described on the basis of the documents; analysing its production process, however, is again likely to require querying stakeholders. In a similar vein, one may assess the effects of reporting by consulting certain proxy indicators for 'use' (for example, a report's download statistics as a measure of readership) yet one will necessarily require qualitative methods such as interviews with stakeholders to understand the bigger picture, including unintended consequences. Whenever primary data (e.g., opinion or user surveys) are not available, researchers commonly resort to asking experts on a set of research questions (Spada & Paulson, 2023). Stakeholders involved in the production of reports may, of course, have their own agendas or (unknown) biases. Each method and proxy indicator has its limitations. This study therefore uses triangulation: combining document analysis, stakeholder interviews, field visits, and secondary data to increase validity and reliability.

In the social sciences, this type of research is referred to as 'mixed methods'. According to reviewers of citizen initiatives, in comparison to qualitative or quantitative studies, a mixed methods approach "provides a better understanding of research problems than either method by itself" yet "a mixed-method evaluation will require more effort of researchers, as it is usually time-consuming, expensive and researchers will need unprecedented access to information" (Igalla et al., 2019, p. 1188). These are important constraints to consider. In the case of the current research project, another important set of constraints concerns languages and locations. The sustainability reporting practices selected as case studies for the four empirical studies are from various countries including five in Europe (France, Germany, Ireland, Switzerland, The Netherlands) and ten in Latin America (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Mexico, Paraguay, Peru and Uruguay). Data collection therefore required field visits in various countries, document reading, and semi-structured interviews with local stakeholders in Dutch, English, French, German, Spanish, and Portuguese.

1.4 Dissertation structure

As outlined in the research questions, this dissertation adopts a two-by-two comparative design, structured around two world regions (Europe and Latin America) and two macro-topics: (i) sustainability reporting as an empirical phenomenon, and (ii) the relationship between sustainability reporting and local government performance, including practices of voluntary benchmarking and formal accountability. To address these four research questions, four empirical studies were conducted. In each region, one study investigates the implementation and perceived effectiveness of urban sustainability reporting initiatives (focusing on trend monitoring), while the other analyses how such reporting mechanisms intersect with local government policies and performance frameworks (focusing on accountability). Together, these studies serve to explore a spectrum of information-based governance ranging from soft, voluntary disclosure to hard, mandatory accountability.

Each empirical study responds to a sub-question that is both descriptive and analytical in nature. Initially, the relevant sustainability reporting practices are traced and characterised in terms of their emergence, governance arrangements, and content. Subsequently, their effects and long-term relevance are assessed, thereby enabling comparative insights into the merit of diverse reporting strategies in distinct socio-political contexts.

While the four empirical studies each address a specific sub-question of the dissertation, they also contain individually formulated research questions. This reflects their original design as stand-alone articles prepared for peer-reviewed journals. As a result of this publication-based structure, the phrasing of research questions or sub-questions within these chapters does not always correspond

exactly to the dissertation-wide sub-questions. Nevertheless, each study contributes a distinct perspective that aligns with the overarching analytical framework and collectively supports the synthesis of insights.

Following this introductory chapter (Chapter 1), which has been structured according to the logic of a research funnel (cf. Section 1.3), Chapters 2 through 5 present the four empirical studies. Each chapter is self-contained but thematically aligned with the others. The final part of the dissertation, Chapter 6, offers a comparative synthesis and discusses the broader implications of the findings. This chapter addresses the overarching research question and reflects on the methodological and conceptual limitations of the study. It also outlines directions for future research and provides policy-relevant recommendations.

In metaphorical terms, the funnel of the introduction is inverted in the final chapter, thereby giving this dissertation the well-known *hourglass* structure that is visualised in Figure 1.8.

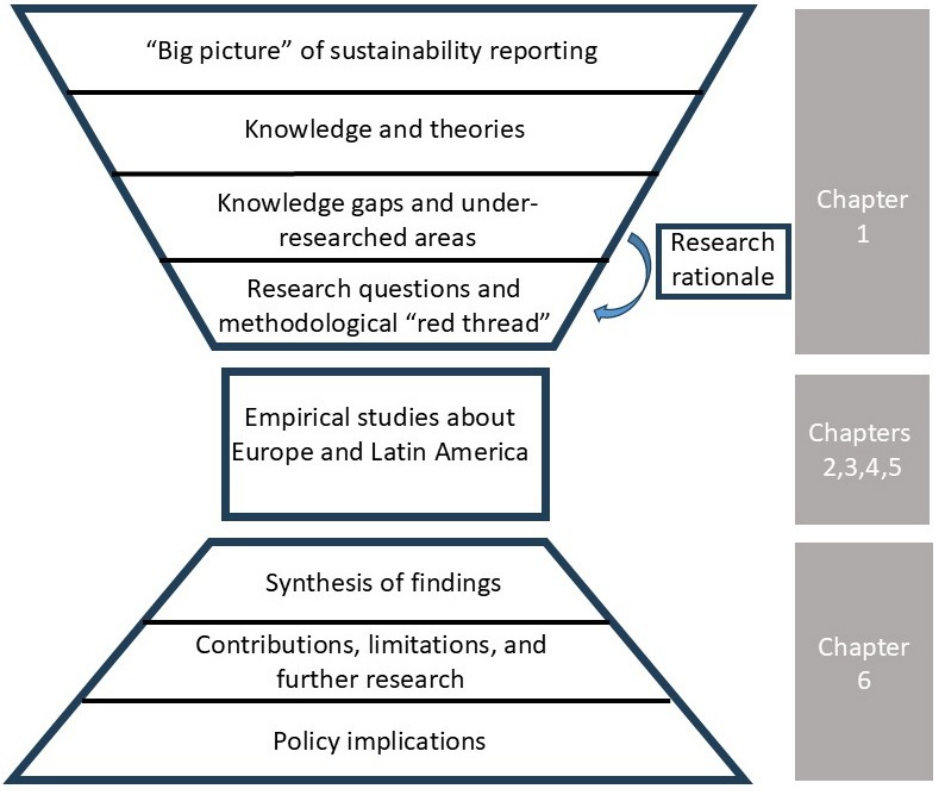


Figure 1.8 *Hourglass structure of this dissertation*

2


Dublin
Amsterdam
Freiburg
Nürnberg
Basel
Zürich
City
sustainability
report

Sustainability reporting for European cities: Lessons from six local government frontrunners in four countries

This chapter is based on the article entitled “Sustainability reporting by local gov-ernments: a magic tool? Lessons on use and usefulness from European pioneers” that was published in *Public Management Review* (http://dx.doi.org/10.1080/14719037.2017.1293149)

2.1 Introduction

> "It's great to have an absolutely rigorous reporting framework. The only problem is they're useless unless somebody uses them". Civil Servant, Dublin City Council (I.16).

Sustainability reporting is on the rise throughout the public sector. International frameworks such as the United Nations 'Sustainable Development Goals' (specifically SDG target 12.6) call for increased reporting by all types of institutions. In the European Union, a recent directive (2014/95/EU) requires all large 'public interest entities' to start disclosing 'non-financial and diversity information'. France recently mandated all municipalities with more than 50,000 inhabitants to periodically produce sustainability reports (Commissariat Général au Développement Durable, 2012), and similar legislation is mooted elsewhere.

Proponents of reporting applaud this trend, with the Global Reporting Initiative (GRI) as the internationally most influential institution. Media coverage can be emphatic too: referring to the positive experience of Amsterdam and other cities, one news article claims that 'a commitment to sustainability reporting is a vital step towards creating vibrant cities' (Ballantine, 2014, p. 4). According to their prefaces, sustainability reports generally serve ambitious objectives targeting multiple audiences. The mayor of the German city of Freiburg, for example, writes that "this sustainability report, presented to the municipal council and the public, serves in conjunctions with the municipal budget as an important management instrument for sustainable urban development" (Stadt Freiburg, 2014, p. 3). In analogy to 'magic concepts' such as accountability and governance that are popular in public management (Pollitt & Hupe, 2011), sustainability reporting is thus often portrayed as a magic, 'jack of all trades' tool simultaneously fostering better policy-making and citizen engagement.

What is the evidence for reporting being an effective, multi-purpose, universally applicable way of promoting sustainability, both inside and outside of local governments? There are conjectural statements about various positive effects (e.g. Lamprinidi and Kubo 2008) but also warnings: In the private and public sector, some critics fear 'accountingisation' where sustainability reports are merely 'an outlet for 'greenwashing' or a source of 'managerialist' information' (Dumay et al., 2010, p. 543) that "may reinforce business-as-usual and greater levels of unsustainability" (Milne & Gray, 2013, p. 13). In a less extreme scenario, reporting may lack or lose its benefits; as evident from this article, in some cities, sustainability reporting was started with enthusiasm but later stopped following what practitioners describe as 'reporting fatigue'.

Surprisingly, academic literature shows little consideration for these real-world phenomena. Most studies explore why organisations decide to become reporters and what kind of information they typically 'disclose', implicitly assuming that

transparency and thus reporting is worth pursuing. In sustainability reporting research, few studies address processes, and too easily the question of 'how' gets a response in terms of a 'why' (Stubbs & Higgins, 2014). Recognising the need for reflection, this journal [Public Management Review] published in 2012 the assertion that "research which simply focuses on enhancing public sector reporting practices without a broader theoretical engagement in the social and organisational context of the public sector is likely to be misguided" (Lodhia et al., 2012, p. 645). Arguably, sustainability reporting risks "merely exacerbating the already overwhelming amount of disclosure provided without adding further insight" (S. Adams & Simnett, 2011, p. 294) if negative outcomes such as 'information overload' (de Villiers et al., 2014) remain ignored.

This study responds to calls for context-sensitive, process and outcome-oriented research by investigating the use of sustainability reporting by local governments from a longitudinal perspective. Building on evaluation research (Weiss et al., 2005), we developed a framework designed to analyse different reporting practices, and explored those developed by six pioneering cities in Europe since 2004. This study thus pursues the following overall research question: How have sustainability reporting practices evolved in pioneering European local governments, and what are their effects?

This study is organised as follows: Section 2 contains a review of research on sustainability reporting by local governments. Section 3 introduces a framework designed to assess the use of sustainability reporting by local governments. Section 4 addresses research methods including case selection criteria. Section 5 presents the results of the comparative analysis of six cases studies. Section 6 concludes with a discussion of the results and the formulation of hypotheses for future research.

2.2 Sustainability reporting by local governments

It has been observed that the concept of sustainability has "saturated the modern world" whereas "sustainability practices for public services have been neglected by scholars and others as a subject of theoretical research and in-depth investigation" (Guthrie et al., 2010, p. 450). External reporting is one core feature – along with the performance measurement and accruals accounting – of many reform processes of the recent decades (Marcuccio & Steccolini, 2009). The focus on reports has also been fuelled by increased attention (from policymakers and researchers) to accountability (Bovens, 2005; Willems & Van Dooren, 2012). In fact, external reporting arrangements are likely to feature in any performance management or public accountability discussion (Downe et al., 2010).

Given the popularity of both sustainability and external reporting, it may come as no surprise that the conceptual joint venture of 'sustainability reporting' has

become influential. In the public sector, however, the phenomenon is not easy to study. To begin with, the term *sustainability reporting* has two key meanings: (i) producing reports yet also (ii) disclosing information. This dual meaning stands at the root of two major lines of research with different conclusions: Firstly, when assessing the prevalence of reports, a common observation is that "the uptake, forms and practice of sustainability reporting among public agencies is still in its infancy compared to the private sector" (Lamprinidi & Kubo, 2008, p. 328); when studying mere disclosure, scholars praise increasing compliance rates (Navarro Galera et al., 2014; Williams et al., 2011).

The search for disclosure – detecting the presence of desired indicators in institutional communications – has become the dominant research paradigm. A recent review of 178 studies concerning private and public sector reporting classified 58% as applying a form of document analysis (Hahn & Kühnen, 2013). Our review of the literature focusing on local governments produced a similar picture (See Table 2.1).

Table 2.1 *Recent studies concerning sustainability reporting by local governments*

Reference	Study object	Data collection	Main results
Marcuccio and Steccolini 2005	12 Italian local authorities	• Interviews	• Growth in reporting is "management fashion" • Reporting can be a tool for appearing innovative and progressive
Guthrie and Farneti 2008	7 Australian public organisations	• Content analysis of disclosures	• Reporting practices are diverse; the annual report is only one medium for disclosures
Marcuccio and Steccolini 2009	15 Italian local authorities	• Analysis of Social Reports	• No standard set of factors can explain differences in disclosure practices • Major driver: seeking of legitimacy
Plawitzki 2010	German municipalities	• Prevalence of reports • Six case studies	• 80 German local governments have produced one or more reports • Reporting serves management functions, sometimes communication
Williams, Wilmshurst, & Clift, 2011	Australian local governments	• Mail survey	• Half of the respondents report on at least one sustainability domain • 40% of non-reporters indicate willingness to report in future
Greco, Sciulli, and D'onza 2012	10 Italian and Australian city councils	• Interviews with managers and accountants	• City councils have a large degree of discretion on what to report

			• Some differences between countries appear related to cultural differences
Goswami and Lodhia 2014	4 Australian city councils	• Content analysis of annual reports	• No city council had stand-alone reports, many cover sustainability issues in annual reports • There is a need for a holistic, contextual reporting framework addressing local issues
Alcaraz-Quiles, et. al, 2014	55 Spanish towns	• Disclosure of GRI indicators on websites • Regression analysis on 13 factors	• Disclosure driven by individual interests of managers and policy makers
Navarro Galera et. al, 2014	33 European local governments	• Content analysis of websites	• Local governments are reasonably transparent • No link between disclosure and (financial) development

In the private sector, sustainability reporting is often attributed to the objective of maintaining a 'social license to operate', and public sector sustainability reporting research also refers to legitimacy-seeking behaviours. Various studies listed in Table 2.1 affirm that local governments generally have institutional and political motives to adopt reporting practices, "mimicking managerial manners" (Marcuccio & Steccolini, 2005). Cities are keen to strengthen their credentials as being 'green' and 'smart' to gain a "competitive edge in the global knowledge economy" (Yigitcanlar & Lönnqvist, 2013).

In the times of 'open data', however, local government disclosure takes place via different media (print or electronically), different documents (e.g. plans, reports, policy papers), at different intervals, and may be a stand-alone activity or part of a larger process. Furthermore, indicators can be used descriptively or with performance-oriented targets and rankings, which has profound management implications (Behn, 2003). Therefore, studying mere disclosure faces ceiling effects and loses analytical power as it eschews the question of organisational use.

The alternative of researching the production of reports – as opposed to information – is challenging too. How many local governments have sustainability reports? This is difficult to answer in the absence of straightforward conceptual boundaries on what constitutes sustainability reporting. While some documents labelled 'sustainability reports' are little more than indicator tables, others are extensive accounts of trends, actions, and plans. Putative sustainability reports may carry idiosyncratic titles such as 'City X – Progress Account'. Standardised frameworks – e.g. the GRI's guidelines that are widely used by companies – have

failed to catch on among local governments whose reports show considerable diversity without recurring to the GRI (Williams et al., 2011). GRI estimates to have information on about half of all reports applying its guidelines (personal communication, 7 April 2015). The GRI's publicly accessible registry (http://database.globalreporting.org) currently lists 450 'public agencies' – mainly public enterprises – and about 50 cities including Melbourne in Australia and Incheon in South Korea. With reporting remaining voluntary in most countries, there are no reliable registries nor estimates of reporters. Presumably the vast majority of the world's local governments has never (consciously) engaged in sustainability reporting yet some 'early adopters' have multi-year experience.

Furthermore, distinguishing 'reporters' from 'non-reporters' becomes even more difficult when organisations forgo distinct reports while rather including sustainability considerations into their general reporting cycle. Advocates of integrated reporting – designated <IR> by the International Integrated Reporting Council (2008) – herald this approach as an effective way to increase the relevance of sustainability information for decision-makers. The aim is to promote 'integrated thinking' and to overcome duplications and 'silo thinking' by integrating information systems of internal and external reporting (Stacchezzini, Melloni, and Lai 2016). Early discussions included the idea for an organisation's integrated report to be a high level overview yet the current IR framework proposes replacing other forms of reporting (de Villiers et al., 2014).

Critics ironically dub IR the "new holy grail" (Milne & Gray, 2013) and talk of 'capture' by the accounting profession (Flower, 2015) yet it is increasingly popular among companies and also influencing the public sector (Bartocci and Picciaia 2013). As with sustainability reporting, most academic studies have explored determinants of IR adoption. Thus focusing on the "icing rather than the cake", "only a few papers attempt to assess the consequences (costs and benefits) of integrated reporting" – Perego et al. (2016, p. 6) therefore argue that quantitative studies based on publicly available data are inadequate and call for more qualitative research. In the words of Mitchell et al. (2008, p. 68), "we must look beyond what is presented in reports, and evaluate the impact on those involved in the process of reporting".

Specifically for local governments, research is required to understand which type of sustainability reporting leads to which type of short-term and long-term benefits and constraints. Moreover, research requires context-sensitivity: A recent meta-analysis of social accountability mechanisms concluded that ostensibly identical tools (e.g. participatory monitoring) can be effective or not depending on how they are embedded in local policies, and thus only "the subnational comparative method can reveal patterns of variation that otherwise would be hidden by homogenizing national averages" (Fox, 2015, p. 356).

2.3 A framework to study reporting practices and outcomes

Researching the local use and usefulness of sustainability reporting requires appropriate models and methods. Since no single theoretical framework is able to cover the range of reporting practices (Marcuccio & Steccolini, 2009), we seek to make a contribution by constructing a theoretical framework that facilitates the understanding of reporting outcomes. For this purpose, we draw on relevant literature and guidelines about reporting contexts, features, processes and outcomes.

Concerning contexts, a first observation is significant diversity: Local governments share the fact of governing limited, subnational geographical areas such as towns or counties yet this takes place in different economies, cultures, environments, legal systems, etc. In the sphere of sustainability management there are thus important debates about the appropriateness of developing standardised as opposed to context-specific approaches and tools (Joss et al., 2015; e.g. Moreno Pires et al., 2014). Concerning sustainability disclosures by local governments, research points to the influence of administrative cultures and data availability (e.g. Alcaraz-Quiles et al., 2014; Krank et al., 2010), making it plausible that reporting practices too are influenced by such macro-contextual differences.

As for reporting features, since the 1990s, the notion of the 'triple bottom line' has been popular and implies the idea that sustainability reports provide information on environmental, social and economic matters (de Villiers et al., 2014). In 2005, the GRI launched sustainability reporting guidelines for the public sector that suggest addressing three information types, namely organisational performance, public policies, and contextual issues. Figure 2.1 illustrates these with examples.

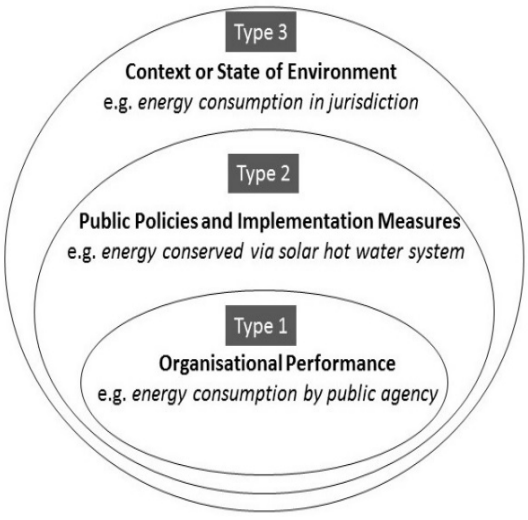


Figure 2.1 *Information types in public sector reports according to GRI (2005)*

For the GRI, "the focus is to provide reporting guidance on the first and second type of information, as the third type of information is often included in other types of reports" (Global Reporting Initiative, 2005, p. 5). To report on

organisational performance rather than on wider, city-level indicators is plausible for many public sector organisations such as utilities or universities. After all, it is "unrealistic to hold agencies accountable for achieving outcomes that are largely affected by forces outside the organisation's control" (Pitts & Fernandez, 2009, p. 403). By the same token, disregarding 'state of the environment' monitoring is unsatisfactory for governments as they are organised along jurisdictional lines and enjoy certain control.

While the GRI framework is about retrospective reporting, failing to address the time dimension beyond comparing a report to the previous one (Lozano & Huisingh, 2011), the IR framework also requires forward-looking projections and targets (Stacchezzini et al., 2016). This implies another information type, labelled 'outlook', and implicitly addresses the question of periodicity. In the private sector, both financial and sustainability reports are usually published annually, strengthening the case for their integration, yet this is no necessity. Many local governments in Germany issue sustainability reports at multi-year intervals (Plawitzki, 2010).

Concerning non-manifest aspects of reporting, practitioners face essential design choices. Who should be involved, what should those involved be doing, and what process should they follow (Mitchell et al., 2008)? For these process questions, normative frameworks offer little guidance. Sustainability reporting as promoted by the GRI recommends the participation and targeting of a wide stakeholder audience, while IR has a narrower focus on providers of financial capital (C. A. Adams, 2015). The latter may be inappropriate for the public sector (Bartocci & Picciaia, 2013); for local governments, key stakeholders relevant for reporting minimally include civil servants, politicians and the public. Another process feature concerns external auditing which according to some authors "should be a permanent element of every sustainability report" (Greiling & Grüb, 2014, p. 220). The evidence base for this assertion is, however, scarce. Most extant research eschews the 'black box' of organisational processes (Perego et al., 2016).

As for outcomes associated with reporting, normative frameworks extol the positive – GRI guidelines do not identify costs but list multiple benefits including enhanced "intra- and inter-departmental coordination", "operating efficiency" and "participation by various stakeholders in decision making and governance" (Global Reporting Initiative, 2005). Integrated Reporting implicitly alludes to potentially negative outcomes when promising to stop "numerous, disconnected and static communications" (International Integrated Reporting Council, 2008, p. 2). IR thus draws attention to information needs and uses, an issue often neglected in transparency agendas. "The more information there is in a report about individual, social, environmental and economic impacts, policies and practices, the greater is the likelihood of information overload for readers" (de Villiers et al., 2014, p. 1045). Evidently, sustainability reporting may have unintended

consequences. With the introduction of reporting, some non-profit organisations apparently experienced that "morality was replaced by the financial bottom line" (Dumay et al., 2010, p. 534).

Some scholars distinguish informative and transformative reporting impacts; the latter is about external, communicative and the former about internal, managerial perspectives (Perego et al., 2016). Since many of the (purported) benefits of sustainability reporting concern organisational issues (see Lamprinidi & Kubo, 2008), it appears expedient to conceptualise these in more detail. In this regard, a framework used in evaluation research that distinguishes three types of information use produces worthwhile insights. According to its main proponents, "instrumental use is presumed to yield decisions of one kind or another. Conceptual use yields ideas and understanding. Political use yields support and justification for action or no action" (Weiss et al., 2005, p. 14). This typology allows the clustering of manifest impacts whilst also paying attention to the "politics of policy-making" (Bauler 2012). In their prefaces to sustainability reports, for example, mayors commonly express the wish to promote accountability, yet beyond such socially valued aims, any publication can also be used as 'ammunition' in political debates (Lyytimäki et al., 2013). Since 'use' connotes intentionality, some researchers suggest also probing for information 'influence', noting that use does not imply influence, and influence does not require conscious use (Lehtonen et al., 2016).

Table 2.2 *Outcomes associated with sustainability reporting*

Perspective	Uses and influence	Associated outcomes, concepts	Typical stakeholders or target groups	Visibility of outcomes
Internal	Conceptual	• **Organisational change** • Knowledge, ideas • Networks • Motivation • Efficiency	• Staff in local government and associated organisations	• Low
	Instrumental	• **Management** • Decision-making • Evidence-informed policy-making	• Elected politicians • Senior civil servants	• Low
External	Political-symbolic	• **Communication** • Agenda-setting • Legitimacy, accountability • Justification for (non)-action	• Citizens • Media • Businesses • Civil society groups	• Very low

Table 2.2 juxtaposes three main uses, influences and outcomes of sustainability reporting with associated stakeholders, acknowledging that this is a simplification of complex matters. There is evidence of (public sector) sustainability reporting leading to organisational learning, facets of improved management, and positively valued communication with outside audiences (e.g. S. Adams & Simnett, 2011).

One might argue that sustainability reports can also foster learning or decision-making among citizens; after all, some mayoral prefaces profess that wish. Prior research on local government reports, however, found negligible citizen uptake (Steccolini, 2004). It has been asserted that "as far as the public is concerned, the publication of performance data in annual reports and government white papers is for the most part equivalent to putting a message in a bottle and throwing it into the sea" (Pollitt 2006, 52).

Concerning external audiences, the agenda-setting influence of reporting is more plausible. Table 2 further addresses the observability of outcomes, which we rated as generally low for internal ones and very low for external ones. This conjectural assessment, informed by the literature (e.g. Sébastien & Bauler, 2013), serves to highlight the complexity of researching the effects of sustainability reporting. Most organisational changes and learning processes are incremental and influences on external stakeholders highly ephemeral, necessitating qualitative research methods (Perego et al., 2016).

In summarising the literature and concepts introduced so far, we posit that three main factors (roughly corresponding to independent variables) influence the effects of sustainability reporting by local governments: (i) context, (ii) reporting features, and (iii) process characteristics. Concerning (ii) features, the three information types proposed by the GRI (*context*, *policies*, *organisational performance*) plus *outlook* derived from IR constitute relevant constructs; for (iii) processes, prior literature lends face validity to the distinction of organisational involvement, political efforts, and dissemination strategies. Finally, adding organisational change, management and communication as outcome-oriented research categories (dependent variables) leads to the assessment framework presented in Figure 2.2.

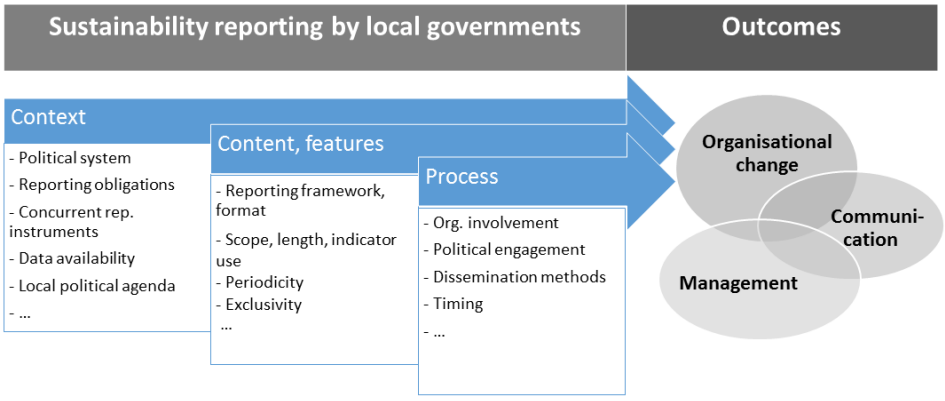


Figure 2.2 *Assessment framework linking reporting factors and effects*

2.4 Research methods

To test the assessment framework, we applied it to real-world practices. In light of evidence that the adoption of reporting is influenced by short-lived fashions (Marcuccio & Steccolini, 2005), we considered it most instructive to study cases where sustainability reporting was developed over longer periods. The purposeful selection of 'early adopters', 'frontrunners' and 'reporting champions' as cases is a well-established methodological choice (C. A. Adams & Frost, 2008; e.g. Bebbington et al., 2009). Our main selection criteria thus were the implementation of sustainability reporting over several years and positive appraisal from researchers, peers or the awarding of public prizes. To ensure a minimum and comparative level of administrative capacities, we chose to focus on European cities with at least 100,000 inhabitants. The six Dutch, German, Irish and Swiss cities presented in Table 2.3 were identified via the literature (see Table 2.1), report registries including the GRI's, and by consulting experts in international organisations including GRI and ICLEI.

Table 2.3 *Case study cities*

Country	City	Inhabitants	Justification for case study selection
Switzerland	**Zurich**	401,000	Multiple reports since 2004
	Basel	195,000	Multiple reports since 2005
Netherlands	**Amsterdam**	830,000	Multiple reports since 2005; use of GRI
Germany	**Nuremberg**	517,000	Multiple reports since 2009; positive evaluation by Plawitzki (2010)
Ireland	**Dublin**	1,270,000	Multiple reports since 2010; use of GRI
Germany	**Freiburg**	220,000	Innovative reporting methodology; use of GRI

To explore the context, history, emergent practices and effects of sustainability reporting in accordance with the theoretical framework (Figure 2) required applying mixed research methods. As virtually all outcome areas including organisational change and agenda-setting are not directly observable (cf. Table 2.2), we applied a document analysis and exploited insights retrieved from key informants. In each city, semi-structured interviews were held with three types of informants: civil servants, elected politicians (mayor or city councillor), and academics or NGO representatives. Adding the GRI, 19 interviews (each lasting 30 to 90 minutes) with 21 informants (numbered as I.1 to I.21) were held during visits and by telephone. All interviews were tape-recorded, transcribed and key statements translated into English for interviews held in German and Dutch. Data were collected and coded according to the assessment framework (Figure 2.2) and the additional operationalisation presented in Table 2.4. For each discernible reporting practice, the first author qualitatively rated four reporting features, namely the comprehensiveness and quality of information concerning context, public policies, organisational performance and outlook. Further qualitative ratings were applied to the aggregate intensity of three process features (organisational involvement, political involvement, and dissemination efforts) and the perceived strength of effects in three outcome clusters (organisational change, management, communication). The rating was omitted when interviewees lacked knowledge, e.g. of processes dating 10 years ago.

Table 2.4 *Operationalisation of research constructs*

Dimension	Subdimension	Coding and rating of evidence for:	Information source
Content	**Context**	• Indicators (e.g. CO_2 emissions) concerning the jurisdiction • where useful: comparisons, benchmarks	• Analysis of published report
	Public policies	• Policies, investments under government control • desirable: analysis of impact of decisions	
	Organisational performance	• Organisational accounts (buildings, staff, etc.) • Performance indicators, e.g. CO_2 emissions of governmental buildings • where useful: benchmarks	
	Outlook	• Upcoming plans and decisions • Desirable: Local agenda and targets	

Processes	**Organisational involvement**	• Involvement of local government staff and other stakeholders	• Interviews
	Political involvement	• Involvement of local government (senior) staff and city councillors	• Interviews • Documents
	Dissemination efforts	• Launching events, media, internet • Timing	• Interviews • Communications
Outcomes	**Organisational change**	• Sustainability-related knowledge gains • Data management capacities • Internal and external collaboration • Staff morale	• Interviews
	Management	• Reference to reports and indicators in government plans • Performance monitoring	• Interviews, communications
	Communication	• References in debates, publications • Media response • Information requests • Emulation of reporting by other organisations	• Interviews, media statistics (if available)

This study design has evident limitations: From six, purposefully selected, diverse yet well-resourced 'pioneers' one cannot generalise findings to hypothetical 'average local governments'. While the theoretical framework assumes the importance of context factors, the small case number did not allow exploring these in detail. The fact that cities were identified by name may have contributed to biased responses from key informants, though this risk was mitigated by triangulation strategies. Experimenting voluntarily with diverse strategies, local governments had little performance pressure, and interviewees expressed much openness to share shortcomings and critical reflections.

2.5 Results from six cities

The six analysed 'early adopters' all initiated sustainability reporting voluntarily. Over the years, each deliberately made different major design choices. This section presents a comparative assessment, starting with observations on features, processes and outcomes before identifying potential causal links.

At the level of reporting features, in four cities we identified one major practice such as tri-annual reports in Nuremberg and annual reports in Dublin. For

Table 2.5 *Comparative assessment of reporting features and perceived outcomes*

		Zurich		Basel	Freiburg
Report features	**Type**	**Multi-year report**	**Biannual report**	**Yearly report**	**Multi-year report**
	Editions	2004, 2008	2014	2012, 2013, 2014, 2015	2005, 2010, 2013
	Framework, format	• Own • printed + PDF	• Own (GRI index) • printed + PDF	• Own • web only	• Own • printed + PDF
	External audit	No	No	No	No
	# of pages (last edition)	58	108	31	67
Content	**Context**	++	0	+	++
	Public Policies	+	++	+	++
	Org. Performance	+	++	+	+
	Outlook	+	+	0	+
Process	**Organisational involvement**	?	++	+	+
	Political involvement	?	+	+	++ Council debate
	Dissemination efforts	++ Events	0	+ Internet	+
Outcomes	**Organisat. change**	+ (initially)	++	0	+ (initially)
	Management	?	++	0	0
	Communication	+ (initially)	0	+	0
Status	**Current situation**	Discontinued	To be continued	To be continued	Discontinued; plan: IR

Amsterdam			Nuremberg	Dublin
Multi-year report	**Section in annual report**	**Multi-year "Agenda"**	**Multi-year report**	**Yearly report**
2005, 2008	Since 2010	2015	2009, 2012	2010,2011, 2012, 2013
• GRI-3 • printed	• Own • PDF	• Own • printed + PDF	• Own • printed + PDF	• Own (GRI index) • PDF only
No	Yes	No	No	No
96	20	74	126	35
0	+	+	++	++
+	+	+	++	++
0	+	++	+	+
+	0	++	+	+
?	0	+	+	+
?	+	++ Council vote	++ Council debate	+
?	0	++ Events, media	++ Events, media	++ Internet
?	0	+	++	++
0	+	+	+	+
?	0	++	++	++ (initially)
Discontinued	To be continued	To be cont. with yearly reports	To be continued	Discontinued

Zurich, we distinguished two phases, as its local government initially published longer, multi-year reports before changing to shorter, annual ones. Amsterdam experimented with various strategies including stand-alone sustainability reports, followed by the integration of sustainability indicators into its general (annual, financial) statements, and the launching of a 'sustainability agenda'. The last two instruments do not represent typical sustainability reports; nonetheless, analysing them appeared vital to understand the evolving system. According to official documents, Amsterdam will relaunch dedicated, annual sustainability reports to "enter into an intensive dialogue with the city" while "the financial statements can then focus on managing on the basis of the results and targets" (City of Amsterdam, 2015, p. 60). In the six cities, we thus identified nine sustainability reporting practices that are summarised in Table 2.5. Evidently, discerning reporting practices is not clear-cut; in every local government, multiple reporting instruments exist in parallel and evolve over time.

Comparing features of stand-alone reports revealed diverse information quantities (e.g, documents with 31 pages in Zurich, 126 in Nuremberg), the absence of external auditing, and the predominance of own, tailor-made formats. Amsterdam (initially), Dublin and Freiburg made loose references to GRI guidelines yet all local governments argued that no existing framework met their needs, prompting the development of their own formats. In the view of several key informants, comparisons and benchmarking were desirable but local relevance an overriding concern. In the words of one civil servant (I.13), "it's good that there are no standardised indicator sets. At most a menu makes sense where local governments can choose which indicators are important for us".

The analysis of content quality showed a mixed picture. Most reports addressed questions of *context*, *public policies*, *organisational performance*, and *outlook* to some degree. In Zurich, switching from multi-year to annual reports brought reduced coverage of context and outlook issues. Freiburg's report stands out since it pays detailed attention to (select) *public policies* and *organisational performance* while lacking city-level outcome indicators, a *context* feature common to most sustainability reports. From one edition to another, reports usually discuss long-term trends through a set of indicators (ranging from 21 in Basel to over 100 in Nuremberg). In addition to such continuity in monitoring, Nuremberg's reports contain changing focus themes (e.g. 'education').

The exploration of reporting processes showed that in most cities, report writing was led by staff units that engaged other departments in indicator selection and the drafting of narratives. In some cities, external stakeholders were also consulted in the design stage (e.g. universities in the case of Zurich, Dublin and Freiburg). Generally, draft reports then underwent a screening process by political decision-makers such as members of the municipal executive. In some cities, finalised reports were discussed in the municipal council, evidencing *political*

involvement. In this context, exclusivity plays a role – Freiburg stressed that its sustainability report, seemingly a stand-alone document, was actually highly integrated into the policy cycle because councillors received it as sole annex to the (biannual) budget. According to a politician (I.2) from Amsterdam:

> If sustainability reporting is separate, we will discuss it separately in the council. When reviewing the annual report, somebody might have a question about sustainability but the discussion will be about finances. Thus, a separate report gets more attention

Interestingly, Basel opted for the contrary when its local government recently decided to discontinue sustainability reporting and merge it with general reporting. According to a Basel political executive (I.5)

> In our system, we have departmental reports but they remain at the bottom of the drawer […]. This is why we decided […] to join cyclical assessments, planning, general reporting and sustainability reporting in a four yearly rhythm. […] Of course, the danger of integration is that the 'flying altitude' rises, with much more general, noncommittal accounts

Contrary to practices elsewhere, Nuremberg's report did not undergo extensive internal vetting. In the words of a civil servant (I.13), "these are the environmental department's analyses. Neither the mayor nor others ever objected. Our critical views are backed up by indicators nobody can simply refute". Interestingly, reporting is thus used politically within the collegiate as part of debates between the economics and environmental departments.

For the *dissemination* of reports, all local governments recently used websites and social media. Usually this involves making reports available for download (with Dublin's not existing in print); only Zurich visualises its data on a dedicated dashboard. Additional dissemination efforts varied significantly, as a civil servant (I.8) from Zurich explained:

> For the first two reports, we did larger events […], a media conference with councillors. We phased that out. Now our main motivation is actually an obligation – we can't just leave the website unattended. In 2013 we wanted to do a public event but it was difficult to get attention. Which is understandable from the media perspective. The news value is not so large.

Regarding outcomes, it appeared that even though none of the local governments explicitly started sustainability reporting with the aim to learn from the writing process, such effects were evident in all cities. Key informants frequently mentioned more inspiration, motivation, cooperation, and improved data management systems triggered by the inter-departmental elaboration of indicator sets and narratives. According to one civil servant (I.13), "it has been a lot of work to

bring all indicators together but when we produced Nuremberg's first sustainability report it had a resounding effect in Germany. We printed 1000 copies which were gone in no time".

Unsurprisingly, there is tentative evidence for a link between organisational involvement and organisational learning. Freiburg officials, for example, mentioned improved morale resulting from extensive staff consultations. However, such relationships appear to be non-linear. Several organisational benefits, e.g. improved data management capacities and collegial contacts, were associated with the elaboration of a first report and not consecutive ones. Concerning outcomes in the sphere of management, stronger effects were observed in Amsterdam and Freiburg. In these cities, there is evidence of reports being actively used by decision-makers. This appears to relate to the presence of two main content factors: targets and politically salient information. In Zurich, for example, politicians showed keen interest in a public perception survey (one of 21 sustainability indicators) as this reflects on voter opinions. However, a civil servant (I.8) also observed:

> Setting targets is political. That needs to be backed up; we cannot do that as municipal administration. Even using a traffic light – we tried to discuss this in the steering group yet realised that this immediately leads to controversial discussions [...]. Our sustainability monitoring system is not the central management instrument of our city governments [...]. All departments have their own key indicators.

As for outcomes in the sphere of communication, most local government lacked insights and relevant data (e.g. media statistics) about the reception of reports by external stakeholders. Only in Amsterdam, Dublin, and Nuremberg, there was anecdotal evidence of active resonance among local audiences such as newspapers, businesses, NGOs or universities. In Dublin, Nuremberg and Freiburg, reporting triggered contact requests from many other cities, nationally and internationally. There is no evidence that any reporting methods were directly emulated in other cities or organisations. Some local governments received public recognition for their reporting, e.g. Nuremberg in the awarding of the German Sustainability Prize.

Concerning negative outcomes, several cities experienced frictions when civil servants perceived report writing as a burden. Critically, in some cities including Dublin and Zurich, key informants spoke of "reporting fatigue"; positive effects associated with initial reports were perceived to wear out in the face of decreasing internal "learning curves" or reduced public interest. In the words of a Dublin civil servant (I.16): "One of the first things about the high frequency is that it ends up being a lot of work. Unfortunately, you end up repeating a lot of things. There's no new data. [...] There's definitely the idea of consultation fatigue". Among the

six local governments studied, four (Amsterdam, Basel, Dublin and Zurich) discontinued or substantially changed their sustainability reporting practices due to dissatisfaction with the approach taken hitherto. This suggests that designing a reporting system with continued use and usefulness is no easy chore.

Among early adopters, Nuremberg's strategy appeared most continuous. Its ostensibly successful strategy, within its particular context, is the elaboration of extensive, low-periodicity reports with a fixed indicator set yet changing focus themes. A Nuremberg politician (I.14) remarked: "I am not a friend of yearly reporting because especially the big issues – air quality, education – don't change that quickly".

At a macro level, Amsterdam's experience and decision to develop several sustainability-oriented planning and reporting instruments in parallel suggests that learning, management and communication – and associated internal and external audiences – require distinct strategies. This interpretation also fits Freiburg's approach: Its highly complex sustainability report – initiated when accruals accounting and performance-oriented budgeting became compulsory for local governments in this part of Germany – implicitly targets councillors, not the public. In the words of a civil servant (I.10),

> We have excellent sectoral reports that go into detail. Our sustainability report can't emulate this. [Its] contribution is to create an overall context [...] to stimulate ideas or to identify trade-offs and goal conflicts. For example, we all want to promote public transport and cycling, that's a declared aim but also requires using space and cutting some trees [...]. We don't have a red or green traffic light but want to show the municipal council its options for action.

Evidently, action-orientation helps to increase salience of reports for specific target groups such as councillors. In this regard, political systems constitute an important contextual factor. As a politician from Dublin (I.17) remarked, "in Ireland, the amount of competencies of local government are fairly small. So one could have a fairly incomplete picture if one only looked at the city council's own activities".

2.6 Discussion and conclusion

In response to our research question – *How have sustainability reporting practices evolved in pioneering European local governments, and what are their effects* – this study covering six cities in four European countries showed that various types of reporting can be valuable for local governments as a learning, management and communication tool. Financial costs can be very limited; some local governments do not even print their reports and rely only on electronic dissemination. There is evidence of organisational benefits, for example concerning

increased staff motivation and data management capacities. This is noteworthy as internal changes were usually no explicit objective. This finding is in line with other studies, e.g. the perception of sustainability reporting as a learning experience in universities (Ceulemans et al., 2015) and of voluntary yet publicised self-assessments stimulating sustainability policies among local governments (Niemann et al., 2017). Organisational outcomes, however, tended to be strongest during the inception of sustainability reporting and to dissipate over time, while (continued) usefulness for management purposes and external communication appeared more difficult to achieve. Various local governments experienced 'reporting fatigue', leading to the discontinuation or radical altering of sustainability reporting practices.

This study's findings tentatively suggest that meeting different information needs of different stakeholders requires smart strategies such as combining extensive, multi-year reports with executive annual updates disseminated in various media. For some local governments studied, especially those producing stand-alone reports, the pursuit of public legitimacy is an explicit objective, corroborating prior studies (Marcuccio and Steccolini 2009). In some cities, this endeavour is relatively successful, with active dissemination efforts producing desired resonance (such as the launching of reports in public events). Various local governments, on the other hand, struggled to maintain public interest over time. This appears to relate to a lack of news value when repeated reports showed unchanged trends on descriptive indicators yet also to a lack of politically salient information.

Most local governments meticulously vet narratives so to accommodate diverging opinions. In one deviant case, the published report is more outspoken and used for 'debates' between executives. The space for such utilisation appears to be influenced by political systems and cultures (cf. Alcaraz-Quiles et al., 2014), with cases in this study showing sub-national diversity in political embeddedness (cf. Fox 2015). Local governments tying sustainability to policy and budgetary cycles aim to inform management, and there is tentative evidence of this being effective. However, targeting internal decision-makers has consequences for the design and writing style of relevant documents – as also asserted elsewhere (Cohen and Karatzimas 2015), sustainability reports geared towards managers are generally not attractive to citizens.

This suggests that sustainability reporting is no 'magic tool' simultaneously fulfilling communication and management functions; instead, attempts to reach all audiences with a single document are doomed to fail, ushering in 'jack-of-all-trades – master of none'. Simply calling for integrated, high frequency, high complexity reporting is misguided as there are trade-offs between conciseness and completeness (cf. Perego et al., 2016). In the words of one NGO representative (I.12) interviewed:

> If you have the ambition of achieving an integrated approach, you'll quickly face unmanageable amounts of data and thick reports nobody reads. Or you're describing in one chapter what you're doing against soil sealing, and in another [...] the shortage of housing, as if the two were not related. Then you have some cities that simply decide to zoom in on focus areas but are rightly challenged too. The crux is solving this tension between comprehensive and focussed.

As a matter of fact, integration can refer to several dimensions including report types, contents and internal processes, and sustainability reporting requires making choices for each. At the same time, many national statistics offices are mounting sophisticated dashboards that individual local governments need not compete with in terms of disclosing macro indicators. The growth of interactive, electronically linked sustainability reporting formats leads to yet more diverse uses by wide-ranging audiences. Facing these developments and choices, there is high demand for guidance that current frameworks (e.g. GRI, IR) do not provide. The GRI's focus on organisational performance without considering territorial outcomes is unsatisfactory for (local) governments. For them, the linking of policies and actions to territorial outcomes constitutes the ultimate management and accountability demand.

Against this backdrop, the framework developed in this study – revolving around the distinction of main contents (context, policies, performance, and outlook), key process considerations (organisational involvement, political involvement, dissemination) and three outcome clusters (organisational change, management, and communication) – is valuable for multiple audiences.

For practitioners, the framework is a point of departure to initiate or redesign sustainability reporting successfully. While the presence of trade-offs and our analysis of a small while diverse set of frontrunners makes searching generalizable 'best practices' inappropriate, this study uncovered many 'good practices' that practitioners elsewhere will find inspiring.

Table 2.6 identifies such practices in relation to our research framework and case studies. The tentative rating of complexity (low / medium / high) serves to indicate ease of replicability. The selection of context indicators and compilation of relevant data, for example, is relatively straightforward and constitutes a recommendable practice for any city initiating sustainably reporting. Political target-setting and impact assessments of governmental decisions, on the other hand, are complex and require more skills and resources.

Table 2.6 *Positive local government reporting practices*

Dimension	Subdimension	Positive practices	Examples	Complexity
Content	**Context**	• Indicator selection based on local priorities, documenting in report annex the relation ("comply or explain") to standardised sets (notably GRI, ISO 37120)	Dublin; Freiburg	low
		• Long time series, discussion of trends • Comparison of local scores to benchmarks (e.g. WHO air quality guidelines) and averages (similar cities)	various	low (provided local data available)
	Public policies	• Identification and discussion of local government competences and spheres of influence	Zurich (2008): matrix of influence	low
		• Clear link (avoiding duplication) between sustainability reporting and other local government planning and reporting systems	various	medium
		• Analysis of choices for and impacts of decisions taken by local gov. in key policy areas	Freiburg	high
	Org. performance	• Long time series; trends • Targets (short + long-term) • Benchmarking with similar cities	various	medium
	Outlook	• Local agenda with short-term and long-term targets and commitments (by local gov. +other stakeholders)	Amsterdam	high
Processes	**Org. involvement**	• Cross-departmental working groups for mutual learning • Minimisation of additional workload (e.g. through collaboration with universities) • Smart periodicity (e.g. multi-year full reports plus shorter annual updates) to minimise fatigue	various	low
	Political involvement	• Discussion, approval of reports in local gov. council	various	low
	Dissem. efforts	• Media + internet events for various target groups • Conscious timing of reporting (release) to maximise attention	various	low
		• Changing focus themes (in reports + accompanying events)	Nuremberg	medium

		• Reports, visualisations and raw data on website	Zurich	medium
Outcomes	**Org. change**	• Self-evaluation (gains, suggestions for improvement) among report contributors and sample of staff	various (incl. Basel, Freiburg)	medium
	Management	• Self-evaluation (uses, suggestions for improvement) among decision-makers		
	Communication	• Tracking of media references, downloads, etc • Solicitation of feedback from citizens / report users on website • Sharing of experiences with other cities		

Regarding outcomes, Table 2.6 contains the implicit recommendation for sustainability reporters to self-evaluate. Simple satisfaction surveys among report contributors and the monitoring of download statistics, for example, produce important feedback at hardly any cost. Researching the utilisation of reports by various audiences is more complex yet may interest a local university, as happened in Basel. For policy-makers – including those contemplating mandatory municipal sustainability reporting as in France, or subsidised voluntary reporting as in parts of Germany (LUBW, 2015) – the framework and evidence produced by this study will give guidance about policy objectives and local government needs.

For researchers, the theoretical framework and major design choices identified in this article – such as periodicity, information types, process arrangements – suggest major lines of enquiry. In terms of sustainability reporting, we posit that the focus on disclosure as main research paradigm has had its day, and that more utilisation- and impact-oriented studies are urgently needed. There is potential for increased mutual learning with the adjacent fields of performance management (see e.g. Navarro Galera et al., 2008) and city-oriented sustainability monitoring (see e.g. Tanguay et al., 2010). Table 2.7 lists a set of hypotheses derived from this explorative study to inform further research.

Table 2.7 *Hypotheses for further research*

Factor	Hypotheses	Relevant case
Context	• In consensus-oriented political systems, the room for sustainability reports to contain political analyses and targets is limited.	Zurich, Basel compared to Nuremberg
	• The longer the term of councils or budgets, the lower the advantages of frequent reporting.	Freiburg

	• Demands for outcome-oriented budgeting create opportunities to use sustainability indicators for management purposes.	Freiburg
	• More limited municipal competences increase the demand for distinct sustainability reports addressing outcomes in the city-at-large.	Dublin
Content, reporting features	• The inclusion of performance indicators and targets is associated with increased instrumental / management use.	Amsterdam, Freiburg
	• Over time, mere trend reporting based on descriptive indicators leads to decreasing interest from internal and external audiences.	Zurich
	• Frequent, high content reporting with limited news value leads to internal "fatigue" and decreased external interest.	Dublin, Zurich
	• Comprehensive, high content reporting integrating financial information is ineffective as external communication tool.	Amsterdam
Process	• Organisational involvement (e.g. collectively defining goals and indicators) increases organisational learning	Zurich, Freiburg, Nuremberg
	• The simultaneous yet exclusive launch of reports with governmental plans and budgets increases management use.	Freiburg

In light of preliminary evidence that sustainability reporting by local governments can be beneficial at limited costs and risks, we conclude by calling for perseverant yet reflective experimentation. As a civil servant from Dublin (I.16) put it:

> You have to say on the first day: We're going to do five years of sustainability reporting and then evaluate. Just to create that expectation. Because it's very hard to do a good job the first time. You have to learn the lesson.

3


Monterrey
Guadalajara
Bogotá
Cuenca
Lima
Belo Horizonte
Rio de Janeiro
São Paulo
Córdoba
Santiago
Buenos Aires
Red Ciudades

Sustainability reporting for Latin American cities: Lessons from 49 civil society initiatives in 10 countries

This chapter is based on the article entitled “How to Sustain Sustainability Monitoring in Cities: Lessons from 49 Community Indicator Initiatives across 10 Latin American Countries” that was published in *Sustainability* (https://doi.org/10.3390/su13095133).

3.1 Introduction

> We made recommendations on adjustments to the city development plans, some of whichwere accepted. However, relations with the municipal administration have been a bit tense due to the poor results they obtained in our citizen perception surveys. This madethem very upset.
>
> (Quote by the representative of an indicator initiative in Colombia; informant CO#3).

To improve a city's liveability and democratic governance, a strategy with much appeal in this digital, data-driven age involves tracking numbers on issues that many citizens and elected officials care about (Kitchin et al., 2015): How is our city or neighbourhood doing (and comparing to others) with regard to, say, air quality standards, quality-of-life, municipal expenditure, and crime rates? Organisations dedicated to compiling and publicly reporting on local wellbeing or sustainability indicators at regular intervals have been called 'urban observatories' or 'community indicators'; in line with other studies (Barrington-Leigh, 2017), we label them 'community indicator initiatives' in this study. Pioneers such as Sustainable Seattle were founded in the 1990s (Holden, 2007) and hundreds of other local projects were active at the turn of the millennium (Moreno Pires et al., 2017).

In Bogotá in 1998, the Colombian capital's chamber of commerce, a reputed foundation, media firm, and university jointly created *Bogotá Cómo Vamos* ("Bogota, how are we doing"in Spanish). This organisation's core mandate is to publicly report on the city's quality-of-life and sustainability on the basis of relevant indicator data obtained from official sourcesand supplemented with their citizen perception surveys. Inspired by this Colombian model, dozens of similar initiatives subsequently appeared across Latin America (Hevia, 2016; Pozzebon & Mailhot, 2012). Many chose equivalent names (*Lima Cómo Vamos*, *Rio Como Vamos*, etc.) or variants often combining city names and plural first-person pronouns—in Brazil, *Nossa São Paulo* ("Our São Paulo"in Portuguese) was created in 2007 and emulated by namesake initiatives in Argentina such as *Nuestra Buenos Aires*.

To exchange experiences and coordinate joint activities, they collectively created the *Latin American Network of Just and Sustainable Cities* in 2011 that comprised, in its heyday, some 60 like-minded initiatives from 10 countries (Hernández Quiñones & Flórez Herrera, 2013).

Across Latin America, 'community indicators' evidently proliferated in the previousdecade, fuelled by substantial investments in time and money by civil society volunteers,private foundations, universities, journalists, and entrepreneurs. In several cities, inter- national donors, such as the Inter-American Development Bank (IDB), co-financed the creation of a "civil society monitoring system" (Inter-American Development Bank, 2016). A quick internet search for key actors

reveals that *Bogotá Cómo Vamos* (cf. www.bogotacomovamos.org (accessed on 1 April 2021)) and Colombian sister organisations continue to be active, just as *Nossa São Paulo* (www.nossasaopaulo.org.br (accessed on 1 April 2021)) and fellow initiatives in Brazil. However, *Rio Como Vamos* and *Nuestra Buenos Aires* no longer exist, the Latin American network became inactive in 2016, and in some countries, none of the original community indicator initiatives survived.

The rise (and fall) of community indicators in different countries triggers two fundamental questions: What do they achieve? Which contextual factors influence their achievements and survival? Various academic studies address the first question and pointto several positive impacts, including gains in public awareness and sustainability-oriented decision-making (Moreno Pires et al., 2017). In the literature on sustainability indicators at large (thus including their application at a national level, in industries, etc.), scholars have also identified the possibility of misuse (e.g., disinformation campaigns) but non-use appears to be the moreprevalent risk (Lyytimäki et al., 2013). Indeed, failure to achieve long-term usage seems to be a common ailment. According to a recent review by Wray, Stevens, and Holden (Wray et al., 2017, p. 10),

> not unlike the story in other realms of voluntary and community work, efforts in community indicatorshave been plagued by the short lifespan of many initiatives. All too often, the cycle is one of a burst of investment of, enthusiasm, dedication, skill, and resources, a hard slog to establish an initial reputation and reporting system, some small triumphs of media, community, and perhaps even political attention, followed by a series of disappointments in efforts to repeat, accelerate, or institutionalise the work, and ultimately by the decline or disappearance of the initiative.

In that view, fizzling out over time seems to be the normal fate of indicator projects; moreover, unrealistic expectations on the side of practitioners appear to be more important than either the local or national context or the specific choices each initiative makes about its activities and organisational set-up. With regards to internal governance, some researchers offer specific recommendations. Extrapolating from a case study in Australia,Davern et al. (2016, p. 571) posit that "all indicator systems should include these best practice principles in their development and operation", which include the prescription to "include a balanced mix of government, business and community representation". The last point is remarkable since, in Latin America, virtually all initiatives operate at an arm's length fromgovernments and have governance arrangements that explicitly exclude elected officials.

Therefore, it is fair to state that there are both theoretical and empirical reasons to map and evaluate community indicator initiatives in Latin America. Theoretically, various assumptions about the effectiveness and 'best practice' of community

indicators can be put to test, in particular, with regards to prescriptions about organisational governance. As this study shows, community indicator initiatives differ widely in the number of stakeholders involved, their degree of cooperation with media firms, reliance on volunteers, indicator choices, dissemination methods, and other characteristics.

Empirically, the Global South is under-researched, and a more representative selection of case studies is repeatedly called for by scholars of sustainability indicators (Ramos, 2019). Latin American initiatives operate in diverse environments, ranging from smaller towns to the world's largest cities, in countries showing differences in terms of public service levels, political violence, and access to information laws. This provides unique opportunities to open the "black box of contextual drivers" (Grandvoinnet et al., 2015). Responding to calls in the literature for comparative, longitudinal approaches, this study, therefore, sought to answer the following research question:

> *Which design and context factors are associated with the influence and long-term viability ofcommunity indicator initiatives in Latin American Cities?*

Three research sub-questions help structure this article: How do city-level community indicator initiatives function in different Latin American countries? What do they perceiveas their objectives, barriers, and achievements? Which contextual factors, in combinationwith organisational strategies, are associated with their success and failure?

Through our theoretical and empirical contributions, we aim to strengthen the globalbody of knowledge on indicator initiatives and to provide insights to practitioners; the latter include civil society activists, donors, and decision-makers involved in the design of national transparency and accountability policies. We further expect that the research frameworks elaborated for this study, including the typology of context and design factors as well as effects, will inform future studies. Our approach and objectives are summarised in Figure 3.1.

Figure 3.1 *Outline of research issues, approach, and objectives*

This study is organised as follows. The next section contains an overview of essentialliterature on community indicators, with a view to guiding the elaboration of the conceptual model applied in this study. This is followed by a section describing the research population and methods and one presenting key results. The final section contains a discussion and conclusion. Further details and raw data are available as Supplementary Materials [on the journal's website].

3.2 Community Indicators in the Context of Transparency and Accountability Initiatives

Community indicators function at the intersection of three issues of high standing onthe global policy agenda and SDG framework (Klopp & Petretta, 2017) — namely, urban sustainability, transparency, and civic participation. As such, they relate to vast and overlapping academic disciplines, including policy studies, governance, communication, and management sciences (Lyytimäki et al., 2013). To contextualise our comparative, longitudinal evaluation of Latin American initiatives, we reviewed the literature for key findings in three areas: activities of sustain-ability indicator projects, their organisational structure, and their achievements, includinglongevity; a transversal concern in each area is the influence of contextual factors.

Context differences are under-researched regarding community indicator initiativesyet are likely to play a major role in their success or failure. Informed by sectoral reviews,such as the one by Wray et al. (2017), and the transparency and accountability literature, such as Grandvoinnet et al. (2015), contextual factors with major relevance for community indicator initiatives include the political–legal

regime, data availability and the presence of alternative, competing transparency and accountability actions at the country and city levels.

With regards to organisational set-up, a joint venture of governmental and non-governmental organisations is often taken for granted (Moreno-Pires & Fidélis, 2012) or even put forward as the "best practice" (Davern et al., 2016). A comparative review of the longevity of 82 indicator initiatives assessed theparticipation of governments and academic institutions and also concluded that "a broadand cross-cutting alliance of stakeholders is the most promising, because it can best ensureongoing demand for the product, collective accountability, [...] a robustness of funding" (Barrington-Leigh, 2017, p. 25). The same study implicitly describes barriers impairing the long-term functioning of indicator initiatives, namely funding and sustained demand for indicator information.An organisation's legitimacy is also mentioned yet rarely studied comparatively; this is nosurprise, as it is highly contextual and hard to operationalise (Holden, 2013).

Research on indicator projects has been dominated for years by studies exploring how they select which indicators. One main distinction concerns bottom-up, participatory, and context-specific versus top-down, expert-led, standardised approaches, with each having advantages and advocates (Fraser et al., 2006; Moreno Pires et al., 2014). While a set of indicators can also be merged into one index, most initiatives forgo this option, as aggregation obscures meaning and transparency (Barrington-Leigh & Escande, 2016). Sustainability monitoring typically involves outcome indicators, often clustered in the environmental, social, and economic domains (Tanguay et al., 2010).

According to a recent global analysis of 67 measurement initiatives, a city's unemployment rate and green urban space are the most widely used indicators (Merino-Saum et al., 2020). For these two examples, indicator initiatives in rich countries can obtain relevant information from official statistics or they may just need to convert raw data into more meaningful indicators, such as "green space per capita". For subjective wellbeing, however, another indicator that is widely used among community initiatives internationally (Barrington-Leigh, 2017; Cox et al., 2010), and pioneers in Latin America (Lora, 2016; Martinez et al., 2021), all data needs to be collected through surveys, representing an expensive undertaking. In many indicator systems, political processes only receive marginal attention (Merino-Saum et al., 2020), though exceptions have been documented, such as a multi-year, voluntary benchmark among Dutch municipalities exclusively tracking sustainability policies (Niemann et al., 2017). Further, in addition to reporting outcome indicators, many initiatives deal with political processes, either through monitoring policy-related indicators—for example, government expenditures—or active community engagement, for example, by organizing round-tables, public events, and so forth.

Figure 3.2 visualises this array through sample activities found in this study and the literature (Moreno Pires et al., 2017). The set is non-exhaustive, but it illustrates the spectrum of activities involving indicator reporting (concerning outcomes in the city and processes or inputs by the local government) and actions for community engagement. The latter may also be conceptualised as targeting generic processes (e.g., via community events) as well as the outcomes of deliberations, such as formalised spatial development plans. This spectrum mirrors that of sustainability reporting by local governments; a study by Niemann and Hoppe (2018) concerning the practices of European pioneers found significant divergence in terms of content along outcome and process dimensions and various degrees of citizen engagement strategies. A number of North American community indicator initiatives also attempt to monitor government performance at large (de Lancer Julnes et al., 2019). On the other hand, process-focused activities mayalso be carried out by other organisations not working on sustainability. From this perspective, community indicator projects with their diverging scope of community engagement can be conceptualised as belonging to the broad field of transparency and accountability initiatives (Grandvoinnet et al., 2015).

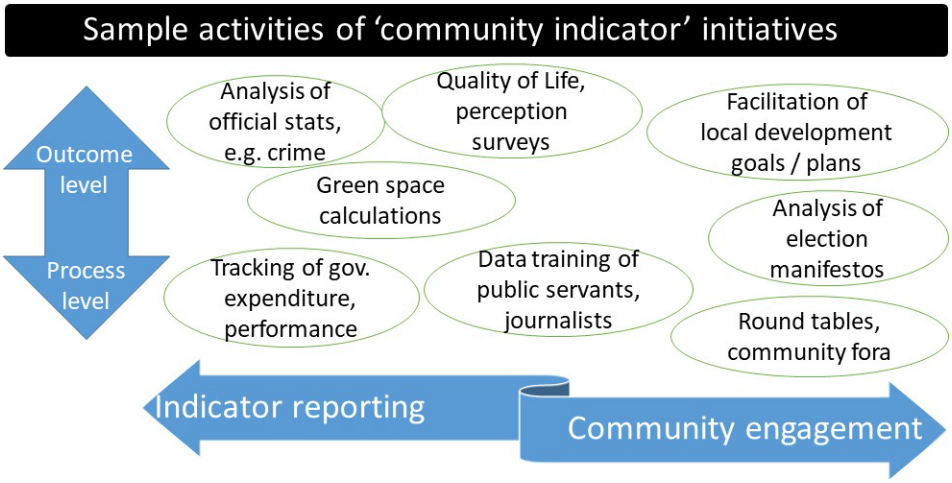


Figure 3.2 *Sample activities of community indicator projects or urban observatories*

Assessments of effectiveness conceptually require taking objectives into account so one can measure progress against actual goals. Naturally, goals differ between indicator projects in different contexts and lifecycles (Joss et al., 2012). Beyond the implicit notion of improving local governance and sustainability, however, initiatives often lack explicitly stated objectives (Wood, 2016). To help assess outcomes in general terms, other studies about sustainability indicators (A. A. Hezri

& Dovers, 2006) distinguished three clusters of use and influence: instrumental (i.e., bearingon decision and policymaking), conceptual (e.g., learning and capacity development among public servants), and political–symbolic (e.g., public discourse). These clusters overlap but help identify intended effects—such as instrumental influence on local government policies—as well as unintended uses and misuses.

A common challenge for research on the effects of 'infomediaries' (Grandvoinnet et al., 2015) and multi- stakeholder actions is identifying causality. As all community indicator projects operate innetworks (of different data providers, users, target audiences, etc.), their societal impact cannot be studied in isolation but requires a systems perspective. The method of process tracing allows for the gathering of high-quality evidence; a study applying it to three American community indicator initiatives found positive impacts regarding agenda-setting and other dimensions (Wood, 2016). Process tracing and similar methods, however, require complex methodological rigour, which cannot easily be applied at a large scale. This explains why large-N comparative studies typically resort to more superficial measures of success, suchas organisational survival or "staying power"(Barrington-Leigh, 2017).

Figure 3.3 summarises the main factors thus identified in this literature review—various contextual factors (at country and city levels) as well as design choices (in particular, about organisational set-up and actions) that have a bearing on what community indicator projects achieve in terms of uses and influence. A further variable of interest is organisational evolution, which also serves as a reminder that none of the other constructs are static but actually co-evolve over time.

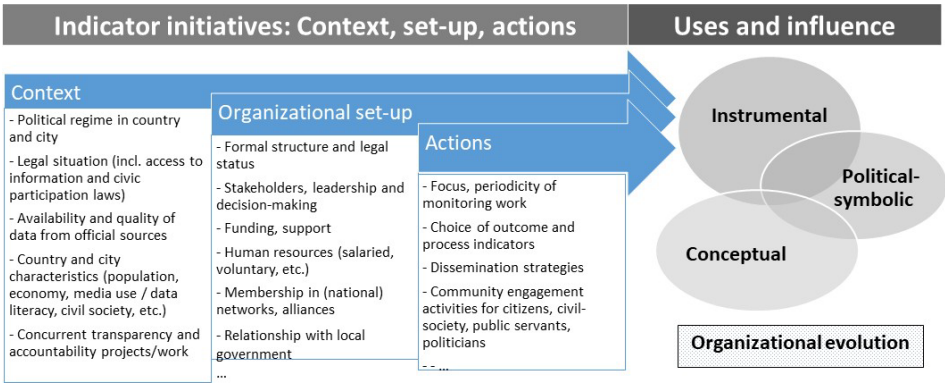


Figure 3.3 *Conceptual framework based on the literature review*

3.3 Population and methods

The world's "battle for sustainability will be won or lost in cities", the United Nations (2012) proclaimed. According to this metaphor, Latin America is an important battlegroundinasmuch as it is the world's most urbanised region, where about 80% of the population lives in cities. In 2011, the Latin American Network for Fair, Democratic, and Sustainable Cities (In Spanish: *Red Latinoamericana por Ciudades y Territorios Justos, Democráticas y Sustentables*) was created by community indicator organisations from across the region. Some,for example, *Bogotá Como Vamos,* had already been working for years with salaried staff. In other countries (e.g., Ecuador), all initiatives were run by volunteers, without funding or legal status, and with just a Facebook page to begin with. In conferences and discussions, ashared purpose was agreed—namely, promoting equitable development and democracy by reporting on sustainability indicators and the local government's management—as well as the stipulation that members are civil society organisations. In terms of joint activities, priority was given to communication (websites, newsletters, etc.) and developing an indicator database for benchmarking.

In the absence of their own financial resources, grants from philanthropic foundations (notably, AVINA) allowed the funding of a part-time communication officer in Lima, a secretary in Mexico, and consultants in Brazil developing a collective indicator database (For joint communication, the network used to run the website www.redciudades.net (accessed on 1 April 2021) that following the end of project funding, has been offline since 2016). Network representatives also decided to support this longitudinal research by sharing the (contact) details of member initiatives and volunteering to be interviewed. Importantly, key informants concurred that the network included virtually all relevant, city-based Latin American initiatives and that no alternative institutional model (such as sustainability reporting by local governments) were prevalentin any country. Therefore, membership, as assessed in 2014, can be considered representative of the population of Latin American community indicator initiatives at the time. Thenetwork counted 65 initiatives on its own website, but verification of individual internet sites showed 8 as actually inactive in 2014. As documented in the online Supplementary File, these initiatives were discarded from this longitudinal study, just as 8 fledgling initiatives only started in 2013 (As also documented in the Supplementary File, at least 9 initiatives have been founded since 2014, including one in Guatemala, as an additional country). Therefore, the main sample consisted of 49 city-based community indicator initiatives started by 2012 and active in 2014.

For research on the practices, experiences, and long-term viability of these Latin American community indicator initiatives, the conceptual model derived from previous studies (cf. Figure 3.3) served as a starting point. For each of its main

dimensions (context, set-up, actions, use and influence, evolution), we identified the key parameters and data collection methods, as presented in Table 3.1. To assess an initiative's context, we relied on 'hard data', such as city population numbers and the UN's Human Development Index at the country level. To explore the political–legal landscape, we further mapped country ratings from four international indices (by the World Bank, Transparency International, Reporters Without Borders, RTI). Such governance indicators are methodologically contested (Høyland et al., 2012; Michener, 2015),unsuited to reliably capture all discontinuities, such as those arising from radically different presidencies in Brazil and Mexico, and not necessarily valid sub-nationally. However,country differences plausibly have repercussions at the city level and comparative indices may be meaningful to show overall trends and country clusters. For further validation and information at the city level, we interviewed and surveyed representatives of indicator initiatives in Spanish (and in Portuguese for Brazil).

Table 3.1 *Operationalisation and data sources*

Construct	Factor	Parameters	Data Source
Context	Country characteristics	• Indices on human development, government effectiveness, corruption, press freedom, right of access to information.	• United Nations, World Bank, Transparency International, etc. • Interviews
	City characteristics	• Population size. • Availability/quality of sustainability data. • Prevalence of political pressure	• Gov. statistics • Survey; interviews
Organisational set-up	Formal structure	• Type of formal registration.	• Internet query; survey
	Stakeholders	• Categories of formal stakeholders; in particular, (local) governments.	• Internet query; survey
	Human resources	• Number of staff/volunteers.	• Survey
Actions	Focus of monitoring/indicators	• Broad set of (sustainability and quality-of-life) indicators and/or government policies.	• Internet query; survey
	Dissemination methods	• Internet channels (website, blog, social media).	• Internet query; survey
	Engagement activities	• Type of activities mentioned.	• Survey
Use and influence	Conceptual (→ capacity development)	• Knowledge, ideas, "data literacy" in public institutions and civil society.	• Survey; interviews

		• Name recognition: among civil servants/technical staff.	
	Instrumental (→ Policy influence)	• Reference to reports and indicators in government plans. • Name recognition among decision-makers. • Evidence of policy changes.	• Survey; interviews
	Political and symbolic (→ public awareness)	• References in debates, publications. • Name recognition among media and the general public.	• Survey; interviews
Evolution	Organisational continuity	• Evidence of activity in the last 6 months (2021).	• Internet query

The main survey contained open and closed questions about each initiative's set-up, objectives, activities, and experiences (including political interference and quality of public data). It was sent electronically in 2014 and responded to by 44 initiatives, a 90% response rate. Most answers were given by a person with high knowledge of internal processes, such as the initiative's director, and included frank statements about sensitive topics, such as relations with local governments. In response to requests for confidentiality, the survey and interview data were anonymised and summarised at the country level without identifying individual cities. Additional interviews were held with 11 representatives of indicator initiatives in 2014, 2019, and 2020.

To assess current activities, the websites of all initiatives were screened in 2021. Thisallowed for reliably mapping whether an initiative publishes indicators (on sustainability,quality-of-life, or government performance) but not their inherently diffuse and less visible community engagement activities; the majority of initiatives lack annual reports. Uses and influences were assessed qualitatively, with estimations (in 2014) of name recognitionamong various target groups serving as one quantifiable proxy indicator.

An initiative's level of continuity in 2021 was assessed through the visibility of activities on the internet. If an initiative's main website no longer existed, alternatives were explored via search engines and social media sites. If the initiative showed substantial activity in any channel (beyond a simple 'like' or retweet) in the last six months, it was classified as 'active'. If no sign of life was found in more than 12 months, the initiative was classified as 'discontinued', and cases in between as 'unclear' (see Supplementary File for details). Average continuity or 'survival rates' were subsequently compared to otherfactors, such as city size and country context. Comparisons based on country averages, however, are only explorative in nature due to the small number of cases involved.

3.4 Results

To structure this section, we present our findings as responses to three research sub-questions. These are descriptive (about the functioning of community indicator initiatives), analytic (about their experiences), and tentatively predictive (exploring the relationship between continuity and context).

Sub-Question (1): How Do City-Level Community Indicator Initiatives Function in DifferentLatin American Countries?

The Latin American community indicator initiatives studied hail from a geographically vast area, as they work in 49 cities located in 10 countries between Mexico and Argentina.See the Supplementary Files for details on names, city population, year of establishment, main activities, and website. Table 3.2, below, summarises, at the country level, the number of initiatives that existed in 2014, with certain characteristics in terms of organisational set-up and actions, and that continue to exist in 2021.

Table 3.2 *Number of initiatives per country in terms of characteristics and staying rate*

		Organisational Set-Up				Actions			Evolution
		Stakeholders (beyond civil society):				Monitoring/Reporting on:			Continuity
Country	Total # of Initiatives (2014)	Academia	Media firms	Business/Private Foundations	Local government	Various sustainability indicators	Quality-of-life/perception surveys	Local government policy	Evidence of activity in the past 6 months (2021)
Argentina	**6**	2	2	4	1	6	2	4	2
Bolivia	**3**	2	2	3	/	3	2	/	1
Brazil	**15**	9	7	9	1	11	5	5	8
Chile	**2**	1	?	?	/	2	2	/	0
Colombia	**10**	10	10	10	/	10	10	1	9
Ecuador	**2**	/	/	/	/	1	2	/	1
Mexico	**5**	3	1	4	1	5	5	3	4
Paraguay	**1**	?	?	?	/	?	1	/	1

Peru	**3**	3	2	3	/	3	3	1	1
Uruguay	**2**	1	/	1	/	1	1	1	0
SUM:	**49**	**31 (63%)**	**24 (50%)**	**34 (69%)**	**3 (6%)**	**42 (86%)**	**33 (67%)**	**15 (31%)**	**27 (55%)**

As evident from this table and the Supplementary File, Brazil (15) and Colombia (10) have had the largest number of initiatives. Some were started in small municipalities of just 20,000 inhabitants but the majority in large cities. The average population size is close to 2 million, and from among the 10 largest Latin American cities, all except for Caracas were at some moment represented in the *Latin American Network for Fair, Democratic and Sustainable Cities*.

In terms of organisational set-up, a large majority of initiatives are distinct civil society organisations. Exceptions include those where the community indicator work is run as another NGO's project, or entirely by volunteers, or as a public–private partnership led by a local government. In fact, local governments are only included as stakeholders in 6% of all initiatives. Significantly more prevalent are media firms (stakeholders in 50% of all initiatives), academic institutions such as universities (63%) and business institutions (e.g., chamber of commerce), and private foundations (69%). Whereas initiatives called *Cómo Vamos* are usually governed by three or four-member institutions, others opted for more fluid internal governance; *Nossa Sao Paolo* auto-identifies as a network with hundreds of stakeholders. In 2014, the number of salaried staff varied from 0 (in fledgling initiatives or those organised purposefully as a 'citizen collective') to 13; the median was 3.

In terms of activities assessed in the study period, about a third of initiatives dedicated explicit attention on their websites to their local government's policies, such as ordinances and compliance with election pledges. Most combined reporting about a set ofsustainability-related indicators (obtained from official sources) with those of quality-of-life surveys implemented by the initiatives themselves. In Colombian cities, such subjective perception surveys are a trademark activity and implemented annually via professional pollsters contracted at a high cost. Other initiatives only have periodic surveys. One Ecuadorian, volunteer-run initiative used sociology students assigned by their universityfor polling. All initiatives also carry out a wide array of outreach activities not amenable to simple quantification. A case in point is capacity building; "the personnel of eight municipalities have been trained to fill out the files to collect the indicators. It's a permanent job", explained informant PE#3. Other examples include organizing public debates with mayoral candidates in the run-up to elections (Cabrera Paredes et al.,

2020) and numerous thematic round-tables with officials and experts about issues such as sustainable transport or child malnutrition. In terms of organisational continuity, the verification of internet sites showed that 55% ofthe 49 initiatives tracked since 2014 were assessed as being active in 2021.

Sub-Question (2): What Do Such Initiatives as Their Objectives, Barriers,and Achievements?

According to the charter of the *Latin American Network for Fair, Democratic and Sustainable Cities* (agreed upon in 2011), member initiatives have five common objectives: (i) to monitor a city's situation in terms of quality-of-life, social justice, democracy, and sustainability; (ii) to promote civic participation; (iii) to monitor and influence public policies; (iv) to promote space for dialogue between civil society, the private sector, and the state; (v) to disseminate information and knowledge for informed civic participation and decision-making processes. In terms of target groups and strategies, the sampled initiatives developed a range of priorities. Whereas virtually all (95% of survey respondents) considered elected officials a main target group, about 60% also targeted researchers and staff in public institutions. As a Peruvian respondent (PE#03) explained, "The work is in stages. We are giving priority to young people through actions and training. The media is getting information and workshops, just as the technical personnel of the institutions, the latter primarily on indicator issues. Social leaders are convened; we sign institutional agreements".

Since (indicator) data represent the lifeblood of initiatives, their availability and reliability are a main concern in some countries. In Bolivia, a respondent (BO#1) explained that "there's very little access to information. Normally we gather it from sources such asinterviews and focal groups". In neighbouring Chile, "the official statistics generally havehigh fidelity" (CH#2). At an aggregate level, 7% of survey respondents in 2014 considereddata obtained from national sources as having no or little reliability, and even 28% thought so of data from local sources. The former, however, are often not usable at the local level; as an Argentine respondent (AR#5) explained, "we were unable to access disaggregated data in our locality".

To explore barriers and bottlenecks, the survey contained questions about the prevalence of difficulties in several areas—out of these, finding suitable staff and media attention was not considered problematic by most respondents. Over 50% of them, however, reported significant difficulties in accessing indicator data and finding funding (for running costs and specific projects). Further, on a survey question about political pressure or interference (in relation to areas of work, research, positive and negative news, and involved personnel), 38% of initiatives indicated suffering from them 'frequently' or 'very frequently'; only a third did not experience pressures. Sometimes this was experienced as a matter of evolving trust—in the words of one Peruvian informant (PE#3), "At the beginning of our activities,there was mistrust of our institution and it was accused of having a political overtone—

this on the part of the municipal authorities and also in some cases of civil society". In other instances, however, maturing initiatives felt increasing heat. A Brazilian informant (BR#5) stated, "We are going through a period of great political pressure from the current municipal public administration".

Achievements are largely intangible and hard to measure. To quantitatively assess to what extent initiatives were effective in achieving name recognition, the 2014 survey asked respondents to estimate which percentages of three target groups knew the initiative's work; the reported averages were 26% for the city's general population, 49% for relevanttechnical staff of public institutions, and 72% for decision-makers such as the mayor and councillors. Regarding the latter, individual responses ranged from 1% to 100% between fledging and mature organisations. It is worth mentioning that some indicator initiativesassess their name recognition via surveys, and also have reliable ways to monitor their interaction with office-bearers if they are invited to present quality-of-life survey results in, say, a municipal council meeting. As a Colombian informant (CO#3) explained, "The percentage of the population was measured in our citizen perception survey [...]. Amongdecision-makers, everyone knows us because we constantly interact with them, either requesting information or at different round-tables".

Regarding other outcomes, key informants reported evidence of effects (in some cases backed up by detailed explanations in annual reports) in various dimensions. In terms ofconceptual use and influence, 77% of survey respondents indicated having contributed to the development of capacities among public institutions. As a Peruvian respondent (PE#03) stated, "We have achieved that information-generating institutions, as well as public institutions, disseminate their information and try to update it [as] municipal public servants are finally understanding the importance of data". Regarding instrumental use and influence, 72% of initiatives (according to the survey) reported having had some—and 30%, even 'very large'—achievements in influencing the design of public policies. According to a Colombian respondent (CO#8), "We made recommendations on the elaboration of the municipal development plans [and] 20% of our recommendations were accepted".

A major achievement for many initiatives—especially in Brazil and Argentina—was also thesuccessful lobby for new bylaws introducing a legal obligation for mayoral candidatesto create action plans with targets for various sustainability indicators, about which the elected mayor then has to report back in public accountability meetings; such bylawsare known as *plan(o) de metas* in Spanish and Portuguese (Romanutti & Cáceres, 2020). By nature, the cluster of political–symbolic uses and influences is the most intangible; one economic way to assess them is through subjective perceptions of key informants and proxy indicators such as media coverage. Over 90% of survey respondents reported that their initiative had contributed to the city's agenda, public discourse, and knowledge. As

an Argentine informant (AR#3) observed, “if by impact we understand that the issues are debated or published, we have had important achievements since in general, the information that weproduce is published by the main media outlets”.

Sub-Question (3): Which Contextual Factors in Combination with Organisational Strategies Are Associated with the Success and Failure of These Initiatives?

The analysis of activity levels showed that 27 of 49 (i.e., 55%) of city initiatives sampled in 2014 were confirmed as active in 2021, but 13 were classified as ‘discontinued’ and 10 as ‘unclear’; the latter category applies to initiatives with a functioning website but noevidence of recent activities (cf. Supplementary File). This leads to a conservative estimateof continuity since some of the initiatives classified as ‘unclear’ continue to work as a less formalised citizen movement with rudimentary social media activity. Interestingly, the cities hosting each of these three groups (active, unclear, and discontinued) showed similar average population numbers. Therefore, city size does not appear to be a predictorof long-term viability. It deserves mentioning, though, that small, rural communities were overrepresented among the initiatives that had stopped by 2014 (cf. Section 3). Chile, Ecuador, Paraguay, and Uruguay each had only one or two initiatives in the initial sample, which implies a limited base for generalisations. In these cases, the average ‘survival rates’ thus need to be interpreted with much caution. Nonetheless, thenumber of initial initiatives per country appears to matter. As shown in Table 3.2, Brazil (53%), Colombia (90%), and Mexico (80%) showed the highest continuity rates. The first two also have active networks at the national level (www.redcomovamos.org (accessed on 1 April 2021) and www.cidadessustentaveis.org.br (accessed on 1 April 2021). (The Colombian network also registered *Cómo Vamos* as a national trademark; elsewhere, the brand is not protected and has been adopted by other actors such as a newspaper in Mexico City). With its 90% survival rate, the Colombian model, including its established brand and tight internal governance (centred around a small set of stakeholders including a media firm, university, and chamber of commerce), is ostensibly successful. According to the network’s website, it now boasts 16 initiatives covering 60% of Colombia’s urban population. A similar institutional *Cómo Vamos* model was successfully emulated in Limabut abandoned in Rio de Janeiro. In several countries, network-oriented models prevailed.As a radical example of a community indicator initiative of the non-institutionalised type, a ‘citizen collective’ in the Ecuadorian city of Cuenca has now run successfully for 10 yearswith neither legal personality nor salaried staff. Both tight and loose, network-oriented models of internal governance have thus been shown to thrive. This is also evident from diverging advice given by key informants about institutional strategies: “Make sure to establish an organisational structure very early in order to receive resources and have

professionals working full time" (BR#9) versus "be an inclusive movement and look for less institutionalised operating models" (BR#2).

To explore the relevance of contextual factors, Table 3.3 shows the number and percentage of surviving initiatives per country juxtaposed to the selected sample of socio- economical and governance indices.

Table 3.3 *Survival rate of community indicators in relation to country context*

		Country Statistics				
	Active (2021) (% compared to 2014)	**Human Development Index (0–1)** United Nations, 2018	**Government Effectiveness(1–100%)** World Bank (WGI), 2019	**Corruption Perception Index (0–100)** Transparency International, 2020	**Press Freedom Index (1–100)** Reporters Without Borders, 2020	**Legal Right to information (1–150)** RTI-Rating.org, 2020
Argentina	2 (33%)	**0.83**	**49**	**45**	**29**	**92**
Bolivia	1 (33%)	**0.70**	**25**	**31**	**35**	**70**
Brazil	8 (53%)	**0.76**	**44**	**37**	**34**	**108**
Chile	0 (0%)	**0.85**	**82**	**67**	**27**	**94**
Colombia	9 (90%)	**0.76**	**56**	**37**	**43**	**102**
Ecuador	1 (50%)	**0.76**	**37**	**38**	**33**	**74**
Mexico	4 (80%)	**0.77**	**46**	**29**	**46**	**136**
Paraguay	1 (100%)	**0.72**	**33**	**28**	**33**	**62**
Peru	1 (33%)	**0.76**	**50**	**37**	**31**	**93**
Uruguay	0 (0%)	**0.81**	**75**	**71**	**16**	**92**

Due to methodological limitations, including the small number of cases and untested validity of country governance assessments, statistical analyses are not meaningful. However, the data suggest the presence of country clusters.Chile and Uruguay are, according to international comparisons (with Table 3.3 showing most recent data available), the region's top performers in terms of government effectiveness, press freedom, and low levels of perceived corruption. Bolivia, Paraguay, and Peru represent the other end of the spectrum, while Argentina holds a middle ground. In five of these six countries, the observed continuity rate is below 50%, whereas it is higher in the four countries occupying average positions in regional comparisons of human development (as measured by the HDI) and governance.

(Paraguay had seen the emergence of five initiatives since 2013 that did not attain organisational continuity. The one established in its capital in 2010 and included in this study remains active.)

At the city level, these observations match the prevalence of obstacles reported by key informants. In Chile and Uruguay, the quality of public data sources and political interference was no major concern. In Bolivia, Peru, and Paraguay, however, access to data was a frequently mentionedproblem. Another dimension on which key informants reported diverging experiences ispolitical interference and intimidation at the city level—these were high in Mexico and Brazil but virtually absent in other countries (According to the RTI, Paraguay, Bolivia, andEcuador have poorly developed access to information laws, whereas Mexico belongs to the world's top performers. At the city level, key informants reported that progressive laws can be helpful yet miss immediate relevance for their work if public institutions do not collect sustainability data).

3.5 Discussion and conclusions

In response to the research question "Which design and context factors are associatedwith the influence and long-term viability of community indicator initiatives in Latin American cities?", this study showed that 49 initiatives located in 10 countries had an average continuity or survival rate of 55% after seven years. This is a conservative estimate since some of the remaining 45% may be active in a different or less institutionalised form.Our finding that many initiatives fizzled out over time complements evidence from other continents (e.g., Moreno Pires et al., 2017).

Underneath the regional averages, we found significant differences at the country level. Some countries studied for this article—notably, Colombia— evidenced high continuity rates of 90%, with further institutional growth in recent years. Importantly, community indicator initiatives struggle in different contexts for different rea-sons. Our data suggest that in higher income and more democratic countries, maintaining institutional funding and sustained (media) attention are common difficulties, while the absence of immediate policy impacts constitutes a main reason for disillusionment among practitioners. In poorer or more violent contexts, however, scarce availability of reliable public data and political intimidation often represent additional existential threats for com-munity indicator projects. This finding aligns with the growing body of literature about the fate of transparency and accountability initiatives (Fox, 2015) and makes a new contribution to community indicator research, which has hitherto been biased in favour of high-income countries, while ignoring developments in the Global South.

Regarding their organisational set-up, virtually all Latin American community indicator initiatives are civil society alliances excluding governmental

stakeholders. This differs from other world regions and challenges 'best practice' recommendations about the inclusion of governments (Davern et al., 2016). Moreover, successful Latin American initiatives show diversity in their organisational set-up; whereas some are governed by a small set of founding members (notably, philanthropic foundations, media firms, and universities), others function as looser networks.

The implementation of citizen perception surveys proved an expensive yet effective way of gaining wide (media) attention. This mirrors similar trends found by scholars elsewhere (Barrington-Leigh, 2017; Rinne et al., 2012). Many Latin American initiatives also offer training in data literacy and indicator use to diverse target groups, such as public servants, journalists, and communityleaders. According to key informants, and corroborated by evidence including representative household surveys, several initiatives also gained significant name recognition in their cities among the general public, technical staff of local governments, and political decision-makers, such as the mayor and councillors. Using their standing, some initiativessuccessfully influenced public policies such as municipal development plans; another achievement in several cities concerned the approval of new legal requirements for local governments to incorporate sustainability considerations and citizen-led accountability fora into their planning and reporting systems. For relatively small organisations with usually less than 10 staff members, these are significant achievements in megacities such as Bogotá, São Paulo, Lima, and La Paz, where numerous pressure groups vie for attention.

This study is exploratory in nature, which implies methodological limitations. Tracking dozens of initiatives over time required selecting a limited set of research constructs, and thereby disregarding the complex nature, evolution, and interaction between a country's and city's context and a networked community indicator initiative. However, to avoid blind spots, we supplemented the deductive analysis of constructs derived from previousresearch with qualitative, inductive feedback from key informants.

More research is needed. Our findings and the conceptual models developed (cf. Figure 3.3, Table 3.1) contain leads for further investigations. To facilitate follow-up research,a number of hypotheses (and the countries serving as relevant cases) were identified for each of the main dimensions of the conceptual model, as listed in Table 3.4. In terms of context, for example, we surmise that the demise of civil society-run community indicatorinitiatives in Chile and Uruguay may partially be attributable to the relatively better quality of government-run indicator systems in these countries (cf. Orellana & Vicuña, 2019). In terms of organisational set-up, partnerships with media firms perceived as unbiased (where available) are hypothesised to foster organisational continuity. Further research on internal leadership and the choice of working as broad networks or via a small group of stakeholders will help elucidate the relative advantages of each strategy and how these play

out over time. Such studies may be informed by findings in the literature that any (performance) indicator system requires continual reconsideration (Pollitt, 2018), and that citizen participation tools shouldbe aligned to a local population's evolving 'sustainability literacy' (M. Cohen et al., 2015).

Table 3.4 *Hypotheses for further research*

Main Dimension	Hypotheses for Future Studies	Relevant case/country in this study
Context	The higher a country's development level and public provision of sustainability indicators, the smaller the niche and perceived added value of community indicator initiatives.	Chile, Uruguay
	In countries with too limited availability of reliable public data, community indicator projects are hard to be maintained.	Bolivia, Paraguay, Peru
	Progressive laws on access to information foster the functioning of community indicator projects.	Mexico, Brazil
	The local presence of philanthropic foundations (and supportive government policies), as well as universities, fosters the successful foundation of community indicator projects.	Various
Organisational set-up	Partnership with media organisations perceived as unbiased help community indicator projects thrive.	Colombia, Argentina
	When relations between governments, NGOs, and media firms are highly politicised, creating joint associations is not viable and work in looser networks more effective.	Ecuador
	Establishing community indicator initiatives in multiple cities (using a joint brand) fosters media attention (also via benchmarking of city data) and institutional stability.	Colombia
Actions	Implementing household surveys on quality-of-life perceptions is expensive yet effective for gaining political and media attention.	Colombia, Brazil, Mexico
	Monitoring citizen satisfaction with mayors via surveys ensures salience but increases the risk of being perceived as partisan.	Colombia, Peru
	Monitoring government performance and compliance with election pledges increases the effectiveness of government work.	Brazil

Regarding activities, we hypothesise (based on findings in various countries) that the implementation of surveys that tap citizens' satisfaction with mayors may bring

attention but also accusations of partisan meddling. Beyond these hypotheses, further research is needed for several other issues, such as the precise nature of community engagement strategies, including cooperation with other urban movements (Friendly, 2017), and various ways of using indicators and indices, as well as relating local monitoring efforts to international frameworks such as theUN's Sustainable Development Goals (Thomas et al., 2020).

We conclude by positing that 20 years of sustainability indicator projects in Latin America offers a wealth of lessons for this and other world regions. Arguably, the 'socio-ecological niche' and potential positive contribution of community indicator projects is globally increasing due to three major trends: (i) continued urbanisation, (ii) increased data availability due to public investments, technological developments, and open data laws and (iii) many increases in socio-economic inequalities, recently exacerbated by the viral pandemic. Sadly, the more unequal a city, the more informative and newsworthy a localised comparison of relevant indicators. In this endeavour, collaboration involving re-searchers, practitioners, and policymakers is key, as is stamina. As an Argentine informant (AR#3) recommended to others wanting to start a community indicator project:

> It is veryimportant to have and guarantee the continuity of the work long-term. Keep in mind that we work with people and that the generation of trust is a central point for the success of the initiative. It is very important to achieve the most heterogeneous participation possibleso that different sectors feel identified and represented.

4

Linking sustainability reporting to local government performance in Europe: Lessons from a Dutch voluntary policy benchmark based on policy indicators

This chapter is based on the article entitled “On the Benefits of Using Process Indicators in Local Sustainability Monitoring: Lessons from a Dutch municipal ranking (1999–2014)” that was published in *Environmental Policy and Governance* (https://doi.org/10.1002/eet.1733).

4.1 Introduction

Local governments are high on the international sustainability agenda; the recently adopted Sustainable Development Goals (SDGs) contain a specific goal (No 11) concerning "inclusive, safe, resilient and sustainable cities". Virtually all contemporary policy frameworks comprise targets and indicators, and ever more ratings and rankings appear with the declared purpose of informing governmental policies and decision-making. The common focus is so-called 'impacts' or 'outcomes', which – in the parlance of performance management – are widely shared, top level desiderata such as clean environments, public safety, and more employment.The implicit assumption is that corresponding indicators (e.g. on air quality, crimes, unemployment rates) helpgovernments make the right decisions and investments ('inputs') to obtain appropriate 'outputs' such as more trees, police officers, and trainings for the jobless.

Crucially, while high-level, outcome information is certainly important for accountability and the modelling ofcausal 'results chains' an effective planning heuristic, the practical use of impact or outcome indicators is fraughtwith difficulties, especially at local level. In the non-profit sector, the 'impact mantra' has been criticised as counterproductive, as it draws precious resources away from services and puts undue emphasis on outcomes for which the causal links are unclear (Ebrahim, 2005). Acknowledging the challenge of relating outcome indicators to day-to-day management, guidance about the SDGs contains the rather general advice that "the distinction between outcomes, outputs, and inputs needs to be handled pragmatically, and the design of goals, targets, and indicators should be guided by approaches that are best suited to mobilise action and ensure accountability" (Sustainable Development Solutions Network, 2015, p. 147).

Intriguingly, most academic overviews of local sustainability indicators (e.g., Ness et al., 2007; Tanguay et al., 2010) are remarkably silent about indicator types. One exception is Dahl's seminal review of *Achievements and gaps of indicators for sustainability*, in which he argues that "environmental, economic and social states are the result of complex processes of development. Policy and management interventions will be more effective if aimedat the process rather than the result, but process indicators have been harder to define" (Dahl, 2012, p. 16).

In existing monitoring initiatives, process aspects receive some consideration since a few international 'city rankings' feature a select number of process-oriented, qualitative indicators. As part of Siemens' *Green City Index*, for example, a jury of experts rates a city's climate change policies (Meijering et al., 2014). However, tools that emphasise process-oriented indicators are a rarity, and many such initiatives have struggled. ICLEI's *Local Evaluation 21* self-assessment tool and the *Reference Framework for European Sustainable Cities* (RFSC), for example, have failed to reach envisioned user numbers. It thus appears that process indicators are not

only "hardto define" (Dahl, 2012) but especially "hard to maintain". In this context, one Dutch ranking system stands out inasmuch as it has a 15-year history. Since 1999, the *Local Sustainability Meter* (henceforth abbreviated as LSM) has called nine times, at 1–4 year intervals, on all Dutch municipalities to assess their sustainability efforts and policies rather than outcomes. In some years, up to 90% of all Dutch municipalities responded by voluntarily filling a questionnaire. This is especially remarkable since in the ensuing, publicly communicated ranking, the winner receivesan award and some municipalities can praise themselves as top performers, whilst just as many knowingly end up at the bottom of the league table. In this study, we thus use the LSM case to explore the merits and limitations of local sustainability measurement tools that emphasise processes and policies. The main research question is '*Under what conditions can local sustainability monitoring initiatives focusing on processes viably function, and what are their benefits and limitations?*'.

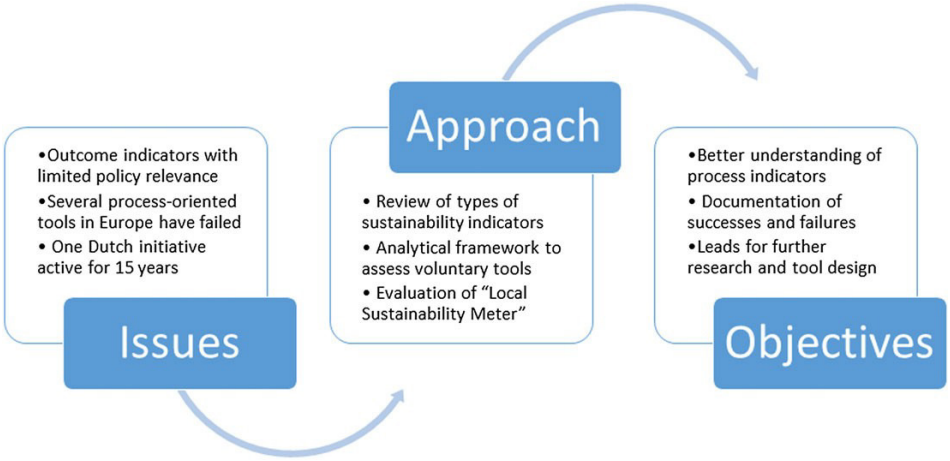


Figure 4.1 *Summary of issues, chosen approach and research objectives*

Our approach and objectives are summarised in Figure 4.1. We set out by reviewing academic literature on process and policy indicators in sustainability ratings and rankings. This is followed by the presentation of an assessment framework. We then describe the research design and methods and subsequently the main results. We discuss their implications and end with a conclusion and suggestions for further research.

4.2 Process-oriented local sustainability monitoring

Since its emergence, the international sustainability agenda has also been a measurement agenda – the Local Agenda 21 (LA 21) movement initiated by the United Nations after the Rio conference (1992) explicitly stressed the need for local

monitoring work. This section summarises key concepts and literature on the focus of such monitoring, on the relation to performance and on their observed effects.

4.2.1 Outcome, process and policy indicators

Monitoring is about identifying trends, threats and progress, and generally built around the tracking of indicators.Outcome indicators have an important role to play in any sustainability assessment, as they allow study of the evolution of key social and environmental concerns across time; many lend themselves to cross-city comparisons (Joss, 2012).

Defining the essence of sustainability and thus choosing sustainability indicators, however, necessarily entails simplification of very complex matters (Turnhout et al., 2007). Even with the advent of ever more automatised data collection, 'big data' analyses and 'open data' gateways, the number of indicators directly informing policies is bound to remain small, especially at the local level. This is because for most sustainability issues, particularly in the socio-economic sphere, goals and norms are relative and contested, and causal mechanisms extremelycomplex – a case in point is poverty, an issue for continuous discussion among policy-makers and researchers about causes, solutions and appropriate measurements. In most policy fields, the ambition of using outcome indicators to steer a city towards sustainable development is hampered by unclear causal attribution, a mismatch between geographical scales (Mori & Yamashita, 2015), time lags and limited data availability (Lyytimäki & Rosenström, 2008).

In light of these challenges, some scholars propose focusing on processes instead. A recent conference report stated that "what is most needed to accelerate and scale up innovation is a process focus, prompting the need for generic, replicable protocols and tools for supporting the design, implementation and assessment of (urban) sustainability initiatives" (Joss et al., 2015, p. 9).

One way of generically assessing process aspects is through policy indicators that probe whether governmentshave a certain type of favoured policy, plan, budget, project or programme. Policy indicators typically use binary scales, e.g. when assessing whether a certain formal policy is present or absent. In the field of human rights, for example, monitoring frameworks commonly track the ratification of international treaties or adoption of laws (OHCHR, 2015). To increase informational value, processes can be decomposed (e.g. formulation of a plan, approval, implementation) or described with explicit demands on the characteristics of components (e.g. the incorporation of certain standards, budgets, approval processes etc.). Another option is to rate policy efforts as 'very strong', 'strong', 'weak' or 'very weak', thus producing cardinal or Likert scales.

However, the more detailed the requirements, the higher the indicator's transaction costs, the less accessible the results become for general audiences, and the less transferable the assessment is across contexts. This explains why early efforts at the times of Agenda 21 to spread the use of standardised 'institutional indicators' failed to gain acceptance. The "share of population that takes part in local Agenda 21 processes" proposed by Spangenberg (2002), for example, defies any simple operationalisation and dodges the question of who should make such measurements. In support of Agenda 21, various European governments issued manuals on how to initiate local sustainability processes and how to assess their quality (e.g. Swiss Federal Office for Spatial Development, 2005) yet these are relatively generic tools relying on the management and research capacities of local governments.

The comparability of processes remains a fundamental challenge. As Dhakal and Imura put it (2003, p. 117), "process indicators are said to help to transfer the knowledge or know-how from one place to another in the formof a combination of qualitative and quantitative indicators. Such process indicators should typically be derived from successful experiences (and best practices) and henceforth should be implemented elsewhere with the necessary modification to suit the local conditions". This suggests that there is a trade-off, with competing demands for standardisation and replicability on the one hand and adaptation for context-sensitivity on the other.

4.2.2 Performance management through self-assessments and rankings

Process indicators can be valuable inasmuch as they inform policy choices and efforts and make the policy processmore transparent. Their use in managing performance, however, is not straightforward. In the field of public management, it is well established that indicators can incentivise desired behaviour but also lead to 'perverse effects' (Hood, 2012). According to Pollitt, tying incentives and sanctions to indicators requires a careful balancing act: "If there is no coupling, or only a very faint connection, then performance targets may not have much effect on behaviour. If, on the other hand, the connection is drawn very tight, so that everybody knows that headswill fall in the event of a missed target, then gaming and cheating behaviours are likely to flourish" (Pollitt, 2013, p. 358). Rankings as one popular performance management tool facilitate comparisons by visualizing relative performance but can also be problematic; frequently criticised are arbitrary weightings of performance indicators,the failure to differentiate inputs and outputs, and comparisons of dissimilar organisations (Tillema, 2010).

Concerning stakeholders external (expert-led) and internal (self-) assessments are the main source of performance information in public governance fields. According to Bovaird and Löffler (2003, p. 326),

> self-assessment is usually more knowledgeable and allows those people to learn the necessary lessons who must later play a key role in improvement

> processes. However, self-assessment is also potentially myopic and self-deluding – and it is lesslikely to be trusted by "outsiders". Conversely, external assessment by 'auditors" or 'inspectors' is more likely tobe independent but in turn is often not trusted by the agency" subjected to assessment. It is believed to exhibit "limited understanding of the context, a tendency to be simplistic or superficial. Moreover it can be a very expensive process" .

To validate process and policy indicators, expert panels are often used, yet this may increase opacity. Some authors thus assert that if "policy indicators are used, binary indicators (determining whether particular types of policies exist in a city) are more transparent than expert-evaluation based policy metrics" (Zhou & Williams, 2013, p. 34). Furthermore, if data are neither derived from statistical, administrative sources (which is the case for most outcome indicators) nor from expert assessments but provided by local stakeholders themselves, one key target audience is directly involved. This may imply biases yet also increase the potential for effective policy learning. In the field of local public service delivery, the use of award schemes with voluntary candidatures is commonplace (Hartley & Downe, 2007). Among Asian cities, winning a sustainability prize has been identified as a powerful motivator (Krank & Wallbaum, 2011). By participating in contests and thus voluntarily undergoing scrutiny, local governments 'signal virtue' (Gugerty, 2009). Competitive awards, however, have evident limitations, in particular their inherent exclusivity. An alternative route designed to reach larger user numbers are certification schemes, e.g. for 'eco-cities' (Joss *et al.*, 2015).

4.2.3 Effects of Sustainability Monitoring Initiatives

Sustainability monitoring in cities generally serves one or more of the following four purposes: Decision-makingand management, advocacy, participation and consensus building, and research and analysis (Parris & Kates, 2003). An American review of 'key indicator systems' concludes that a fully operational set of measures takes timeto develop and requires broad involvement of society and substantial resource commitments, while benefits can include (1) more informed policy choices, (2) a better educated citizenry and (3) greater civic engagement (United States Government Accountability Office, 2011).

Case studies have produced some evidence of positive 'soft' impacts on capacity building, social learning and improved communication between stakeholders (Reed et al., 2006); effects on advocacy, learning and consensus-building are notoriously hard to measure. In recent years, thegovernance and politics of policy indicators has received more attention (Sébastien & Bauler, 2013). Some apparent disappointment felt by practitioners regarding a lack of 'impacts' is arguably due to naïve expectations: "In policymaking environments, a linear thinking of direct, instrumental policy use of indicators is simply not acknowledging the complexity of decision-making, and the inherent discursiveness of policy-making. Not acknowledging

the rich- ness of indicator uses leads to much of the unnecessary frustration at the level of indicator developers, who mightdesperately seek for direct signs of policy impacts of their work" (Bauler, 2012, p. 40).

Self-assessments and process-oriented tools may promote learning and ownership at reduced costs (Joas et al., 2014) but raise questions of potential bias and generally face an uphill battle against low participation among the target group (Garzillo et al., 2014). The rank growth and 'explosion of indicators' (Tanguay et al., 2010) has given rise to many flash-in-the-pan initiatives that vie for attention before fizzling out. For thedesigners and users of monitoring tools, better knowledge of their functioning and effects is therefore highly relevant.

4.3 Towards an analytical framework

How can we distinguish and evaluate local sustainability monitoring tools? Most common are initiatives that periodically monitor quantitative, outcome-oriented indicators obtained from official sources (e.g. national statistical offices), as these are relatively cheap to run. In this process, any local initiative makes various choices in the selection of indicators, their use and interpretation, the frequency of assessment, or the involvement of stakeholders. There are thus no neat classifications, but potentially uncountable configurations. Nonetheless, certain combinations are more prevalent and are thus captured in classifications of monitoring systems and related tools such as process guides (Jensen & Elle, 2007). One framework recently proposed by Grönholm and Berrini (2014) categorises sustainability assessment systems based on the following six criteria: (1) type of evaluation, (2) target population, (3) reach, (4) evaluation area, (5) required efforts and (6) outcome. The first criterion explicitly refers to the distinction between outcome- and process-oriented tools, which tends to be overlooked by other inventories (e.g., Ness et al., 2007; Shen et al., 2011; Waas et al., 2014).

For quantitative, outcome-oriented indicator projects, certain quality standards such as the BellagioSTAMP principles (Pintér et al., 2012) provide guidance that can also be used as starting point for an analytical framework; the same applies for the normative evaluation model proposed by Ramos and Caeiro (2010). For studies on the effects of sustainability monitoring systems, frameworks developed to analyse the utilisation of evaluation and research findings prove useful. The basic distinction of instrumental, conceptual and political or symbolic uses advocated by Weiss is frequently applied (Hezri & Dovers, 2006). These three types of use are inter-related: "Instrumental use is presumed to yield decisions of one kind or another. Conceptual use yields ideas and understanding. Political use yields support and justification for action or no action. Process use tells how evaluation's influence arose" (Weiss et al., 2005, p. 14). Unlike typical outcome-oriented indicator systems, for policy- oriented monitoring that relies on voluntary self-assessments, uptake by the target group constitutes an essential first

criterion of success. Figure 4.2 visualises the analytical framework thus developed for this study. Based on previous research (Grönholm & Berrini, 2014), we hypothesised that various factors including tool and context characteristics explain participation or non-participation. Participation is expected to lead to various instrumental, conceptual and political uses and influences and to strengthen the case for the tool's continuity by influencing the willingness of stakeholders to cover running costs. These causal mechanisms are not clear-cut, however; conceivably a local government can also learn from a tool by using it anonymously, and conversely it might participate without benefitting in any substantial way.

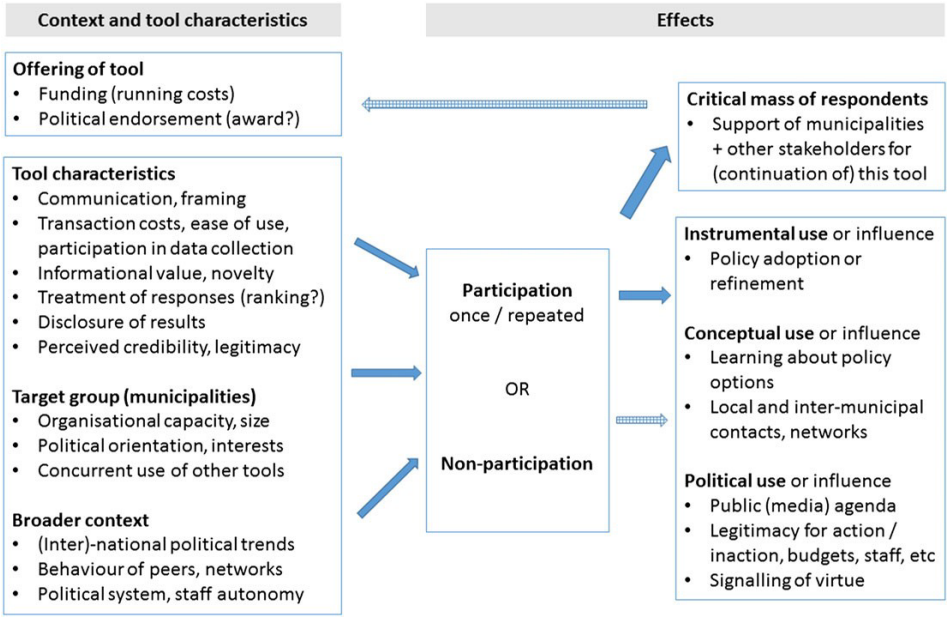


Figure 4.2 *Analytical framework for the evaluation of voluntary monitoring tools*

4.4 Research methodology

This section outlines the empirical methodology of this study, detailing the selection of cases and subsequently ofresearch methods.

Case Selection: When searching for local sustainability monitoring tools based on process indicators, four European tools attract attention – *Local Sustainability Meter, Local Evaluation 21, Baromètre du Developpement Durable* and the *Reference Framework for European Sustainable Cities* –presented in chronological order of their first appearance.

The *Local Sustainability Meter* (LSM) was launched in the Netherlands in 1999. During specific assessment periods – at multi-year intervals – it invites all Dutch municipalities to self-assess their sustainability policies and efforts, producing a ranking and giving an award to highest scorers. In the course of the past 15 years, over 90%of all Dutch municipalities have used this tool at some point.

Local Evaluation 21 (LE21) was designed as "a fully automated self-evaluation tool, simple for the end-user to use, designed for widespread use and offering a fast quantitative evaluation of local governments' engagement in local sustainable development processes" (Grönholm, 2014, p. 58). The organisation ICLEI launched LE21 in 2004, offering it free of charge to all European local governments. Available in 20 European languages, LE21 reportedly received positive appraisal from those who used it. However, in the course of several years only some 150 local governments from across Europe made use of the tool (Grönholm & Berrini, 2014), and it was discontinued.

In France, the LSM's basic methodology was emulated and developed as *Baromètre du Developpement Durable*. BDD employs the same approach of inviting local governments (free of charge in certain regions) to self-assess their sustainability efforts, disclosing responses in public. In contrast to the LSM, it neither has specific assessment periods nor awards, but is open for continuous use. According to its website, the BDD has been used by 314 (about 1%) of French communes so far.

The *Reference Framework for European Sustainable Cities* (RFSC) has been described as "a process-oriented attempt at applying EU sustainability policies" (Grönholm & Berrini, 2014). Released in early 2013, it is a free web-based tool, also available in various languages, offering methodological support by providing an overview of possible actions towards organizing sustainable urban development. Local governments, with small and medium-sized cities being the main target group, are enabled to enter data and link up with peers. However, the EuropeanCommission's financial support for the tool was suspended in late 2014 due to the low uptake, which had remained less than 100 European cities. Observers argue that "the RFSC, deliberately conceived as a non-binding and open web-based tool, lacked legitimizing powers that, in return, has had impact on its capacity to persuade and motivate decision-makers to make full use of it" (Becker, 2015, p. 106).

Table 4.1 summarises key features of LSM, LE21, BDD and RFSC. A quick comparison indicates that each of theseprocess-oriented sustainability tools is different in some regard; for example, the LSM is the only one working discontinuously and offering rankings and awards at multi-year intervals. Further, amongst these four tools, the LSM is 'positively deviant' – it is the only instrument reaching significant participation rates whilst remaining in use for several years, and thus a unique case to assess the viability of process-oriented monitoring. This warrants exploring it in more detail through an evaluative case study.

Table 4.1 *Key features of four policy-based monitoring tools*

Tool	Assessment method, timing	Products for participants / public	Costs and incentives	Coverage / Usage
Local Sustainability Meter (LSM)	Questionnaire with approx. 100 items Data entry by local gov.: for specific calls (every 1-4 years)	Rankings, award ceremonies Scores and rankings disclosed on www.duurzaamheidsmeter.nl	No charges Winners: Awards (endorsed by Ministry of Env.)	Launched in 1999 Uptake: up to 90% of Dutch municipalities
Local Evaluation 21 (LE21)	Questionnaire with approx. 50 items Data entry by local gov.: Any time Optional: input from local stakeholders	Automatised evaluation report (www.localevaluation21.org now defunct) No public disclosure of any data	No charges No incentives	Launched in 2004, limited uptake among European local governments (>1%), discontinued
Baromètre du Develop-pement Durable (BDD)	Questionnaire with approx. 110 items Data entry by local gov.: Any time Obligatory validation by local stakeholders	Scores disclosed on www.barometredudeveloppementdurable.org	No charges in 3 regions; elsewhere need to pay No incentives	Launched in 2008 Current uptake: approx. 1% of French communes
Reference Framework for European Sustainable Cities (RFSC)	Questionnaire / online toolbox Data entry by local gov.: Any time For continuous use (policies, peer exchange, etc.)	Automatised feedback No public disclosure of any data Participants named on www.rfsc.eu (idea: positive label)	No charges No incentives	Launched in 2013, across Europe limited uptake (>1% of eligible local gov's), only 4 Dutch participants

The case study research draws on data compiled from available documents, websites, datasets of LSM editions, and interviews. Table 4.2 lists main data sources. Using semi-structured questionnaires, six key informants involved in the design and management of the LSM were interviewed (two initiators of the LSM, three consecutive coordinators ofthe LSM office and one expert involved in the re-design process of the monitoring tool). Interviews addressed thedesign process of LSM, data collection, usability issues, effects of LSM participation, and strengths and weaknesses of the tool. Interviews lasting between 30 and 90 min were conducted face to face or by telephone. In addition, between 2009 and 2015 multiple LSM advisory group meetings were visited by the researchers. Transcripts andcollected text documents were used to construct a case history. The analytical framework presented in Figure 2 was used to analyse data. Data were coded in order to address information needs arising from the analytical frame- work. Based on transcripts, key interview statements were translated from Dutch to English. This article contains illustrative statements from three informants designated Respondent 1 (a former LSM coordinator), Respondent2 (a technical expert) and Respondent 3 (another former LSM coordinator).

Table 4.2 *Main data sources*

Stakeholders	Main data sources
LSM project organisation	• Website with database covering all LSM editions (www.duurzaamheidsmeter.nl) • Interviews with key informants
Internal stakeholders	• Minutes of meetings of advisory group (containing representatives from Ministries, universities, etc.) • Interviews with key informants • Municipal websites (e.g. www.nijmegen.nl)
External audiences	• Websites of sustainability platforms (e.g. www.duurzamegemeente.nl) • News reports (e.g. www.rijksoverheid.nl/actueel/nieuws/2009/12/24/resultaten-lokale-duurzaamheidsmeter-2009)

4.5 Results

4.5.1 The evolution of the Local Sustainability Meter (LSM)

Following the Rio Conference in 1992, the Netherlands was seen as an environmental pioneer in Europe (Coenen, 2008), eager to adopt the principles of Local Agenda 21. In 1999, the National Committee for International Collaboration and Sustainable Development (in Dutch, NCDO) and the network of Centres for International Cooperation(COS) developed a tool called the Local Sustainability Mirror with the purpose of raising awareness on how localgovernments could formulate policies related to international development. In 2002, COS took sole charge,

changed the name of the tool to Local Sustainability Meter, and shifted the tool's conceptual focus – against the wish of NCDO – to sustainable development in Dutch municipalities. COS later dissolved and FairBusiness, a spin-off company, took charge of the LSM.

Since its inception, running the LSM has been carried out with project-based funding in the range of 30 000–50 000 Euro per LSM edition. The Ministry of Environment provided financial support for the initial design and data collection until 2010; for the last edition, main contributions came from other governmental programmes. Each 'edition' of the LSM can be visualised as taking place in a cycle with various steps: developing a questionnaire, inviting all municipalities to participate, evaluating responses, disseminating results and deciding on a future edition. These activities incur the types of cost listed in Table 4.3.

Table 4.3 *Main activities and costs of the LSM*

	Phase	**Main activity, input**	**Budget share**
❶	Development	Fundraising, conceptual work, stakeholder consultations	Medium
❷	Self-Assessment, basic communication	Creating website, disseminating standard invitations (letters, press, social media)	Low
❷.❶	Active marketing	Time to make personal phone calls	Medium – high
❸	Evaluation	Calculations, evaluation of requests for additional scores	Medium
❹	Dissemination, awards	Results on website, media reports, ceremony	Low
❺	Decision on future	Conceptual work, consultations with stakeholders	Low
	Average cost (funding received) per edition:		**30-50,000 €**

Some of the running costs were constant, such as the services of a helpdesk and maintenance of a dedicated website. The amount of effort invested into actively marketing the LSM fluctuated per edition: with more project funds, more staff time was invested in active marketing. In the words of Respondent 1:

> An important lesson is that 'measurability'– even with regards to quality, policy intentions, and implementation – is easier than I had thought. That personal contact, artisanal and painstaking work, is more cost-effective than glossy folders and broad mailings. I was surprised about the impact of simply making intensive phone calls during two weeks.

According to one interviewee, for the LSM's longevity is has been vital to pursue strategic opportunities and to incorporate trendy topics of current interest, such as corporate social responsibility (2004), sustainable procurement (2007), and more recently climate change mitigation. In 2013, the National Energy Agreement for Sustainable Growth opened up a window of opportunity as this influential platform (agreed between the national and local governments and various societal stakeholders) pledged a continuation of the LSM. However, some parties advocate a merger with two other monitoring tools, namely the Climate Monitor and Sustainability Balance (Zoeteman et al., 2015), that rely on outcome indicators. In this context, some LSM protagonists fear that the future instrument – tentatively labelled 'governance monitor' for the LSM component – will become predominantly outcome oriented, losing attention to process issues and participatory data collection.

4.5.2 Tool characteristics

In this section, main characteristics of the LSM are addressed. This concerns items and indicators, data collection,disclosure and rankings.

Questionnaire design and indicator selection

For the development of its initial indicators, the LSM drew inspiration from governmental and NGO campaigns. The Association of Dutch Municipalities (VNG) and NGOs participated actively, and a panel with 20 local governments tested draft questionnaires. The LSM clustered indicator items under topics ('energy', 'water and nature' etc) before adopting the 'people, planet and profit' trichotomy in 2010. Questionnaire items relate to potential, generally desirable policy actions – for example, the creation of sustainability-related staff positions in the municipality, the request of certain national subsidies, the adoption of climate policies or sustainable procurement plans, or the application of emerging environmental standards in housing projects.

Questionnaire items have varied over the years to maintain relevance – Table 4.4 illustrates the distribution of items per theme and edition. The LSM also featured some one-off topics such as 'sustainable construction'(2000 and 2001) and the 'sustainable canteen' (2008). According to Respondent 2, "dynamic, changing questionsin response to changes in policy options is just necessary in order to keep the questionnaire relevant. Being climate-neutral didn't exist on the agenda 10 years ago". Nonetheless, many items – e.g. on gender or citizen participation – were maintained for years. All items came with the answer categories 'yes'

and 'no', and were assigned weights (1, 2 or 3) that were used for the multiplication of scores. Recent versions also allowed municipalities toapply for up to three additional points per domain for efforts not directly covered by the questions. LSM project staff assessed such 'reward claims' for completeness and merit.

Most LSM items addressed (goal-oriented) policies (e.g. "does your municipality have an integrated plan for the reduction of greenhouse gases?") and not policy outcomes. Less than 5% of items referred to measurable targets. The 2013 LSM edition, for instance, contained one item assessing whether the municipality tracks female participation rates in various projects (policy without threshold) and another one asking whether at least 40% of municipal management posts are held by women (an actual policy goal).

Table 4.4 *Number of questionnaire items per theme*

Thematic clusters	**2002**	**2003**	**2004**	**2006**	**2008**	**2009/12**	**2013**
Social domain, citizen participation	16	16	29	29	29	"People": 27	13
Global – local cooperation	25	25					17
Institutional embedding							15
Climate	27	27	43	43	43	"Planet": 30	20
Water and nature	14	22					17
Spatial planning	19	19					
Waste and pollution							13
Corporate Social Responsibility			30	30	30	"Profit": 37	20
Sustainable Procurement				78	78		20
Sustainable Procurement – canteen				23	23		
Mobility							10

Data collection

The LSM targeted all Dutch municipalities (537 in 2000, 408 in 2013 following various mergers). At the inception,NCDO and COS made heavy use of network contacts (in particular local environmental NGOs) in order to requestthe participation of local governments. Initially, the LSM required municipalities to produce responses in interaction with local citizens' organisations. In later years, LSM turned into an online questionnaire to be filled in for each municipality by its civil servants. Data origin was not fully transparent. In one municipality responses may have derived from a civil servant, whereas in another they came from the 'secretary of the responsible public official', and ina third from a team of multiple actors. Efforts needed to respond to LSM questions were considered time consuming the first time, as respondents had to retrieve information from various municipal departments and local companies or citizens' initiatives.

There is slight evidence that there were validity issues concerning data collection (as reported by technical staff engaged in LSM survey compilation). This concerns forms of data manipulation, with respondents eager to answer 'yes' on questions allowing some scope for interpretation. Moreover, in individual cases a municipal council reportedly blocked the publication of data provided by municipal civil servants. On the other hand, remarkably and perhaps unexpectedly, some municipalities have participated continuously despite repeatedly low rankings. Interviewees indicate that this relates to the autonomy enjoyed by civil servants in the Dutch political context, where local sustainabilityofficers might state their opinions disinterestedly or use low scores to obtain political support in favour of their portfolio.

Disclosure and rankings

An openly accessible website contained all responses. Scores were summed up per domain ('people', 'planet' and 'profit') and in total and used to rank all participating municipalities based on domain-specific and overall scores. Winners as well as losers were thus pointed out, with information publicly available. Since 1999, each edition has been accompanied by an award (the Sustainability Shield) for the municipality with the highest score. Awards were given to the mayors of the winning municipalities in official ceremonies that were supported by national ministers, attracting media attention.

The LSM's questionnaire, website and other tools (e.g. social media) used plain language and contained visual information in various formats including maps. Figure 4.3 shows a sample web presentation from the 2013 edition. The website presents detailed results but deliberately did no downloadable data reports, because municipalities were anxious that consultancy firms would use these to barrage them with unsolicited service proposals.

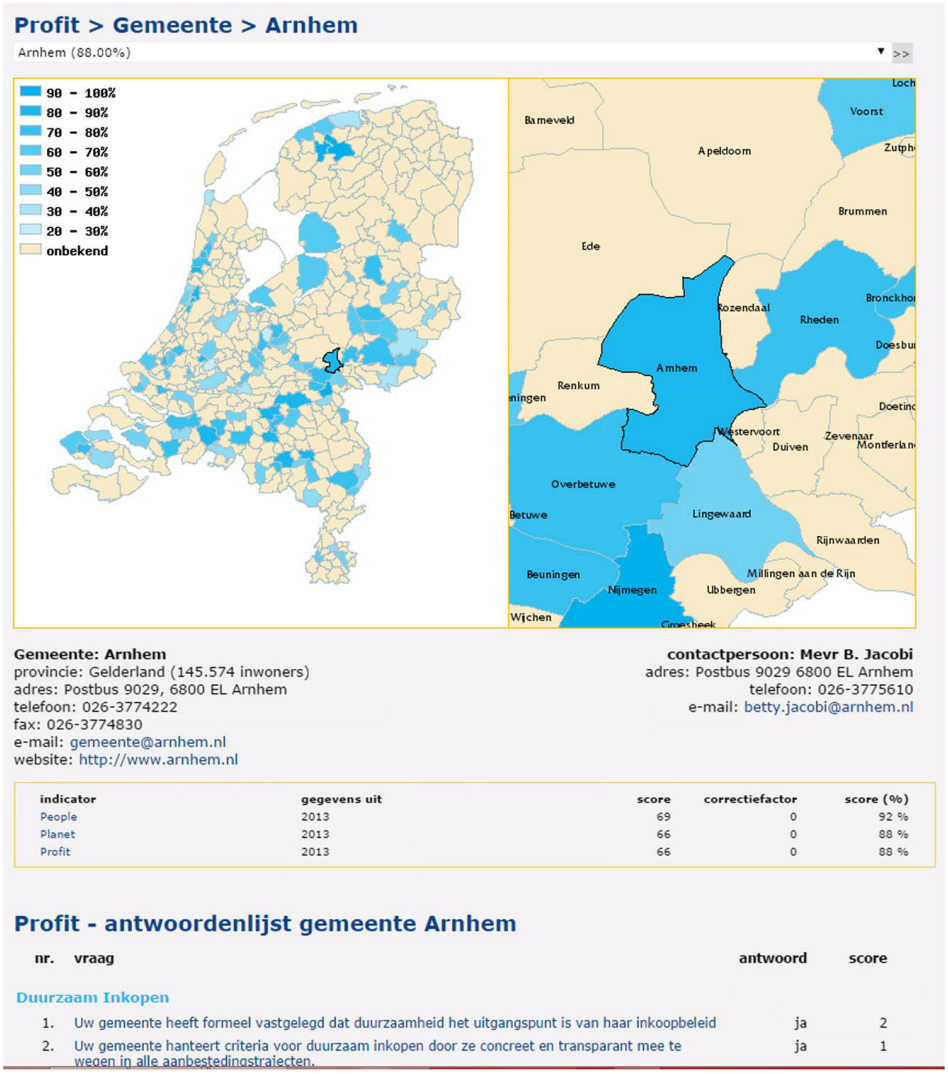


Figure 4.3 *Online presentation of the City of Arnhem's responses (2013)*

Participation Rates

As awareness raising among local governments was considered one of the LSM's objectives, response rates were considered a key indicator of the tool's legitimacy. Figure 4.4 presents a graph of the participation rate of Dutch local governments in the LSM over the period of 1999–2013.

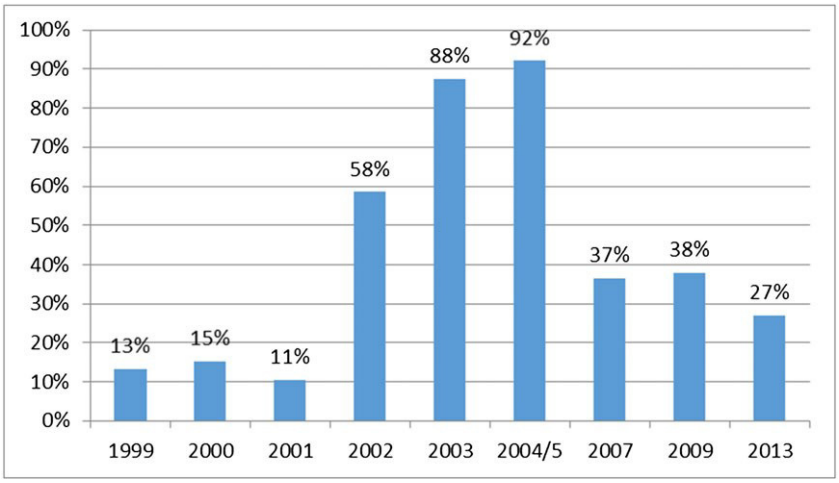


Figure 4.4 *Percentage of Dutch municipalities participating in LSM editions*

The chart shows that during the first three years under NCDO (1999–2001) participation by municipalities was rather low (13%, 15%, 11%). In this period, the LSM office was struggling to obtain responses. When concerns grew about consultation fatigue, in 2002 the LSM team – now led by COS, which had taken over fromNCDO – decided to lower the number of questionnaire items. In addition, a new marketing strategy was deployed, in which LSM team members would personally contact local stakeholders by telephone (especially local environmental group members) using their professional networks. As a result, response rates increased to 58% in 2002, to 88% in 2003 and to 92% in the 2004/5 edition (the latter edition however compiled responses overa two-year period and was therefore not comparable to others). After 2005 – when FairBusiness took over from COS – it was decided to hold bi-annual LSM editions. Response rates, however, declined, from the 2002– 2004/5 period: 37% in 2007 and 38% in 2009. By 2013 the response rate had declined even further to 27%. According to interviewees, endorsement by reputed institutions – notably the national association of municipalities – legitimised the LSM, while the emergence of competing tools – two other instruments besides the LSM have also started to publish rankings of the Netherlands's 'most sustainable municipality' – have contributed to lower LSM participation rates. An examination of response data reveals that large-sized local governments are over-represented; presumably this relates to large local governments having more capacity available than their smaller counterparts (Hoppe and Coenen, 2011).

Instrumental Use and Influence

The LSM has triggered the adoption of sustainable development policies among municipalities. The public accessibility of LSM results allows local governments

to benchmark their own performance against that of others. One interviewee reported that municipalities (at the level of both councillors and public servants) used the LSM 'league tables' to watch evolving performance of peers and to adopt more policy measures before the data collection period was closed formally, thus evidencing the competition effect intended by the LSM. Moreover, 'runners-up' have reportedly also contacted forerunners, asking for advice on what to do to catch up. In the words of Respondent 1, "Some municipalities (perhaps 10–20) have applied the LSM for policy proofing and have also taken on the competitive aspect in their policies, also for reporting instrument to the municipal council. Groningen and a few Frisian municipalities and Tilburg and Breda are examples of this". Interestingly, the LSM questionnaire appears to be used outside the formal assessment period, when incentives do not apply (e.g. by Tubbergen and Dinkelland in 2014).Reported motivations are the wish of a public official, councillor or civil servant to benchmark the performance of their municipality.

Conceptual Use and Influence

One interviewee called the start of the LSM an "institutional and conceptual phase" during which NGO staff learneda lot about municipal practices. According to interviewees, the initial requirement of joint responding by municipalities and local civil society organisations was commonly applied and strengthened sustainability-oriented coalitions as intended. The involvement of NGOs decreased during the 2000s yet throughout the LSM's history there have been cases (e.g. in the City of Wageningen) where local sustainability platforms took the LSM responses of their municipalities as starting point to elaborate a local sustainability assessment report. Respondent 1 recalls:

> Factually the picture produced is most reliable if the LSM is used as intended: in collaboration, in openness, filling the questionnaire or at least going through it together with civil society and stimulating discussions over a few questions. We did that a lot. Over time that decreased and made the instrument a little bit less reliable because transparency decreased.

Responding to the LSM implies reading about local policy options – the questionnaire begins per domain (people, planet, profit) with a written introduction of key terms and policy frameworks. In particular, the succinct referencing of national guidelines has been appreciated by municipalities and helped them in policy-making. Thequestions on procurement, for example, contained explanations on the extent to which European Union and Dutchlaws allow for the inclusion of sustainability criteria, thus bringing practical advice to civil servants. Moreover, the possibility to access peers responses opens up contacts. According to Respondent 2, "There has also been cross- fertilisation between frontrunners. City T. thought that sustainable economic policies were not really feasible but saw that city G. had positively responded to that question. So a phone call was made to find out how they

do it. And you see that this sort of questions also appear in council meetings. So it does contribute to agenda-setting".

Political Use and Influence

The award ceremony for the municipality with the highest score created positive media attention, especially in the early years. Top-scoring municipalities have made explicit political use of positive rankings in their communications. In the words of Respondent 2, "the LSM certainly had influence, and certainly was used as frame of reference during election times". In one case, a municipality reportedly stopped participating after winning the award since it concluded that it had nothing to gain any longer.

Political positioning also played a role within institutions – according to Respondent 3, some civil servants "use the instrument to seek attention for their own role within the municipality". Externally, the LSM has increased the availability of information on government actions to citizens, and this arguably contributed to creating a certain transparency of the democratic process. This approach has inspired tool-makers in other contexts to establish sustainability monitoring systems of their own. For instance, the LSM's model was emulated by Dutch Water Boards, the Dutch provincial authorities and French activists, who initiated the BDD.

4.6 Discussion

What can these findings tell us about the viability and effectiveness of process-oriented monitoring tools? This case study suggests that the LSM's favourable reception has been due to a combination of factors. Some 'enabling conditions' appear to be evident, particularly

- the availability of support (in particular project funding),
- the creation of a recognised brand (endorsement from network partners including respected institutions) and
- the dissemination of policy information in ways that municipal officials consider helpful (e.g. in notes appended to questionnaire items).

These findings about the LSM match observations made in other studies (e.g. Joas et al., 2014) about the importance of institutional support. However, viability over time also requires careful 'balancing acts' (Pollitt, 2013), where the challenge is not maximizing desirable attributes but dealing effectively with competing demands. The LSM hashad to consider tensions including

- choosing the 'right' incentives for participation and performance (ranking, awards) without triggering non-participation and reasons for participants to drop out,
- the advantage of comparability (demanding frequent assessments on identical indicators) in evolving contexts(demanding changing indicators),

- continuity whilst avoiding 'consultation fatigue' of respondents and
- scientific demands (validity of items, context-specific information) vis-à-vis transaction costs and participatorymethods (self-assessment, standardisation).

In dealing with these tensions, LSM managed to create a positively competitive atmosphere that incentivises municipalities to participate and to perform while avoiding the (complete) dropout of low performers. By relyingon self-assessments through questionnaires, it created a model with low transaction costs while tapping the powerof transparency to deter manipulations. This configuration is to some extent context dependent. In the Dutch system, all municipalities have identical legal competencies and civil servants enjoy relative autonomy; in more heterogeneous systems and adversarial political cultures, the LSM approach may play out differently and not necessarily solve the problems of low response experienced by other tools (e.g. LE21, BDD and RFSC). To test this hypothesisrequires comparative research on process-oriented tools in various settings.

The LSM case suggests various lessons on the advantages and limitations of policy-based indicators. These tendto come with neat links to appropriate politico-administrative scales, increasing relevance for target groups (civil servants and councillors). This can also increase democratic accountability, even though it does not necessarily foster citizen engagement, since policy indicators are inherently abstract. Moreover, policy indicators can facilitate relatively holistic assessments. Conventional output indicators – e.g. 'number of trees planted' that the InternationalStandardisation Organisation proposes as a city sustainability indicator (ISO, 2014) – may provide wrong incentives, whereas the (hypothetical) query for an 'integral air quality programme' can trigger the selection of strategies comprising various outputs.

Entrusting local stakeholders to ensure 'integral policies' also helps to account for contextual differences between territories. Furthermore, the public dissemination of information on peers that affirm having developed such policies was shown to stimulate learning. Arguably, knowledge management can be yet more effective if the assessment incentivises municipalities to share and connect (e.g. by uploading policy documents onto the tool's website), but this is a balancing act too; the compulsory sharing of supporting documents raises thresholds for voluntary participation, and virtual networks only thrive with a critical mass of members. This mirrors observations made about the STAR rating system offered to North American cities, which skilfully combines outcome and process indicators; it is resource intensive and requires cities to pay for certification yet gaining popularity as participants benefit from legitimacy and learning (Elgert, 2016).

A potentially more problematic aspect of policy indicators is their privileging of regulatory and standardised approaches to sustainable development that do not pay attention to trade-offs, contextual differences and uncertainties. The public

management literature (in 't Veld, 2010) refers in this context to 'policy accumulation', i.e. the tendency to respond to policy failure with ever more policies at the expense of experimenting with self- governance and local alternatives. Ultimately, policy indicators may incentivise the production of well-phrased policies printed on glossy paper. On the other hand, many policy recommendations implicit in LSM items concerned multi-stakeholder participation, thus strengthening governance beyond municipal competencies. It is plausible that policy-based indicators are more effective when a policy area is less mature – in countries "where local sustainability processes are still in the early stages, a focus on concrete activities is needed in order to capture progress" (Garzillo et al., 2014, p. 117). Presumably, this is a reason why policy-based tools were more prevalent in previous decades and why the response rate of the LSM has declined. Arguably, the term 'Sustainability Meter' is to some extent a misnomer – the LSM is less about measurement than about incentivizing local governments, and as such it is not an alternative but a complement to outcome-oriented monitoring.

4.7 Conclusion

This article explored as the central research question 'under what conditions can local sustainability monitoring initiatives focusing on processes viably function, and what are their benefits and limitations?'. We applied this question to the case study of the Local Sustainability Meter (LSM), a monitoring tool offered to municipalities in the Netherlands. This study showed that an impressive number of municipalities have used the LSM voluntarily. There is evidence that in many cases the tool has had an impact in terms of instrumental, conceptual and political utilisation. Some participating municipalities reported learning – in a competitive yet collegial spirit – from others and improving their (policy) performance; in other cases, civil servants and civil society groups used the LSM for self-evaluation and strengthened sustainability-oriented coalitions and local governance arrangements. Overall, the combination of voluntary, transparent self-assessments at multi-year intervals with public awards has proved to be an effective – and economic – way of disseminating sustainability policies among municipalities. An apparent limitation of this approach is the inability to reach all municipalities, and the lack of context-sensitive policy recommendations. Generalizing from this case, it appears that, under the conditions exploited by the LSM (i.e. the political context of Dutch municipalities) and the design choices made by this tool (self-assessments on changing questionnaires with disclosure and awards), process indicators have been put to effective, long-term use. To advance knowledge on the transferability of this approach, however, more research is needed. One important research question concerns the link between sustainability policies, governance efforts and actual outcomes; to minimise selection biases, one could research a random sample of municipalities. Other fruitful research

topics include a comparative analysis of different monitoring approaches; in response to differing "local socio-political opportunity structures" (Holden, 2013). Various authors have speculated about the importance of cultural factors (Hood, 2012). For the promotion of local sustainability policies, the distinction between outcome and process orientation deserves more attention from researchers and practitioners.

5

Linking sustainability reporting to local government performance in Latin America: Lessons from mayoral accountability obligations

This chapter is based on the article entitled “Should mayors be accountable for election promises? Effects of compulsory goal setting and reporting requirements on sustainability governance in four Latin American cities” that was published in *Frontiers in Sustainable Cities* (https://doi.org/10.3389/frsc.2025.1450933).

5.1 Introduction

In international policy discussions, two ideas have become mainstream in recent decades: the world needs sustainable urban governance, and sustainable urban governance needs goals and indicators (Ferreira da Cruz et al., 2019). Since the 1990s, then promoted by the UN's 'Agenda 21', many cities in Europe, North America, and other world regions have witnessed the emergence of sustainability monitoring initiatives run by municipalities or civil society organisations (Sharifi, 2020; Wray et al., 2017). Initially, a bottom-up approach of selecting indicators based on local priorities was more prevalent; in recent years, the 'top-down' variant of referring to standardised frameworks (such as the Sustainable Development Goals (SDGs) or other indicator sets including ISO's, 2014) has become more common.

In these three decades of goals gaining a central role in urban sustainability governance, one constant in most countries has been the reliance on aspirational goals, the non-committal monitoring of progress, and voluntary reviews by interested parties. Some of these voluntary efforts have had positive effects by influencing a city's and country's sustainability discourse, data availability, and institutional capacities (Lehtonen et al., 2016; Ortiz-Moya & Reggiani, 2023). This is similar to the SDG framework itself, which is also designed as a non-binding agreement (Aust & Plessis, 2018; Biermann et al., 2017) that brought discursive benefits (Biermann et al., 2022) but shows bleak prospects in terms of actual goal achievement (Wu et al., 2023; Zeng et al., 2020).

The trend towards voluntary, non-binding goal-setting arrangements in (inter) national sustainability governance contrasts with an alternative development that also originated in the 1990s and gained prominence in Latin America. In 1994, Colombia introduced legal obligations for mayors to be held accountable for election promises and goal achievement in office; non-compliance or underperformance may trigger recalls. As depicted in Figure 5.1, this accountability mechanism addresses the electoral-administrative cycle by requiring candidates to write explicit manifestos, elected mayors to formalise plans with goals, and mayors to report on goal achievement.

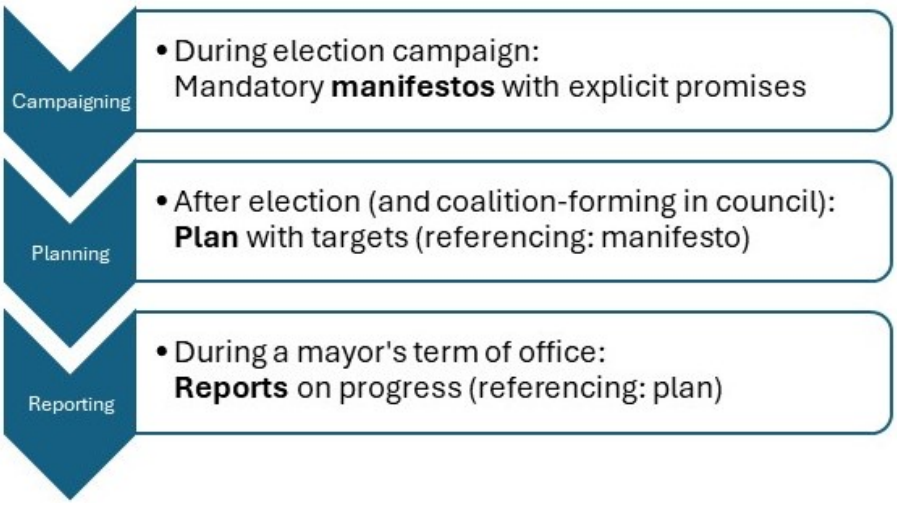


Figure 5.1 *Main phases and components of novel performance accountability for mayors*

According to its proponents (Bravo, 2019; Programa Cidades Sustentáveis, 2020), these new accountability rules should have two major clusters of benefits: One is increased government performance resulting from more programmatic policymaking, less corruption, and less clientelist politics; the other is better functioning (local) democracy resulting from more citizen participation and trust. The combined effects are expected to lead to a more effective 'sustainability governance' at city level. To what extent does this innovation (belonging to "Latin America's experimentalist vocation" (Pogrebinschi, 2018) yet hardly known in other parts of the world), deliver on these promises?

The extant literature contains few answers. Internationally, there are studies about SDG localisation and urban governance (e.g., Morita et al., 2020; Valencia et al., 2019) and some, mainly US-based findings about the relationship between community indicators and government performance systems (Ammons & Madej, 2017; de Lancer Julnes et al., 2019; Greenwood, 2008). Numerous scholars call for more case studies about urban sustainability governance, particularly concerning reporting (Bexell & Jönsson, 2019), experiences in the Global South (Bell & Morse, 2018b; Meuleman & Niestroy, 2015), and longitudinal research (Giles-Corti et al., 2016). In the absence of systematic documentation that would allow a comprehensive mapping or analysis of practices in various Latin American countries, we opted for a comparative, longitudinal case study on the four cities of Bogotá, Córdoba, Guadalajara, and São Paulo. To structure our research and guide the selection of methods, we developed the following two research questions:

1. What are the formal characteristics of novel performance accountability mechanisms that have evolved among pioneering local governments in Colombia, Brazil, Argentina, and Mexico?

2. What are the positive and negative effects of these mechanisms in relation to different design features and contextual factors according to a multi-year evaluation?

Our study's purpose is threefold: (i) to produce an empirical understanding of some performance accountability arrangements for city mayors in Latin America, (ii) to help develop a replicable evaluation framework and research agenda, and (iii) to contribute lessons for policymakers in the field of urban sustainability governance (including the SDGs).

The remainder of this article is structured as follows: Based on insights from the academic literature, we first present an assessment framework designed to study the characteristics and effects of the emerging accountability mechanisms. The second section details the selection of cases and methods. In the results section, we introduce the case study cities and provide evidence on the above research questions. In the discussion, we reflect on study's academic merit and make suggestion for further research. The conclusion contains succinct answers to the overall research question.

5.2 Assessment framework

In line with good practice recommendations to be transparent about the analytical choices made in developing one's assessment tools (Bovens et al., 2008, p. 234) we developed an assessment framework that aligns with our research questions.

Our first research question—addressing the formal characteristics of the new accountability mechanisms—is mainly descriptive and methodologically straightforward. It can be explored via data found in documents and factual descriptions by local observers who are familiar with the long-term evolution of a city's system. Since the research object concerns accountability rules, a useful conceptual tool also applied in this study is the 'accountability cube' (Brandsma & Schillemans, 2012). This model proposes analysing a given accountability mechanism in terms of three dimensions: How much accountability-relevant information is provided (little vs. much, applicable in our case to mayoral plans and reports), the level of discussion triggered (intensive vs. non-intensive, applicable in our case to exchanges between mayors and other stakeholders including the public), and the amount of consequences (few vs. many, applicable e.g. to the threat of recall).

Our second, evaluative research question (on the intended and unintended effects of an accountability mechanism) is both conceptually and methodologically more

challenging. In terms of effects, it is legitimate to assess whether the intended benefits such as 'less clientelist policymaking' are observed. However, this is conceptually not sufficient as all governance innovations are likely to have intended as well as unintended consequences (Dahler-Larsen, 2013; Lehtonen et al., 2016). Therefore, assessing the relative value of a new mechanism requires probing for diverse effects in various domains. Moreover, research on public accountability and performance management has shown that well-intended innovations can have effects that most observers would classify as negative, e.g., when the publication of performance indicators leads to organisational risk aversion, rigidity, and strategic behaviour known as the 'gaming' of targets (Bovens, 2005; De Bruijn, 2002; Pollitt, 2006).

To capture this variety of potential effects—positive and negative, intended and unintended—we build on an assessment framework used elsewhere (e.g., Hezri and Dovers 2006; Lehtonen et al. 2016; Niemann and Hoppe 2018) that distinguishes three main clusters: instrumental, conceptual, and political-symbolic aspects. Instrumental effects would be evident if the accountability mechanism led to more efficient, evidence-informed governmental action (intended and positive) or the gaming of targets (unintended and negative). Typical stakeholders for this type of 'use and influence' of goals and indicators are politicians and senior civil servants. The second cluster concerns conceptual effects as potentially evident from organisational learning and intra-governmental collaboration (positive) or the emergence of a blame culture (negative). Typically affected are staff of government or civil society organisations. The third cluster concerns political, symbolic, and discursive effects such as more attention to sustainability or citizen trust (positive) or the loss of trust and alienation between the government and citizens. The latter, unintended phenomenon has been described as the "tyranny of light" by Tsoukas (1997). Symbolic 'greenwashing' (cf., Velte, 2022) also belongs to this category. Table 5.1 summarises effects and main stakeholders of the accountability mechanisms as expected on the basis of prior research.

Methodologically, the challenge is that most potential effects are intangible. It is thus difficult to observe them empirically and also hard to attribute them causally to any given intervention. On the matter of citizen trust, for example, population surveys could contain relevant data, yet even if they existed at the city level (which is often not the case), one cannot interpret any changes in trust as evidence of the accountability mechanism being benign or harmful. Ultimately, numerous confounding factors at the national level may have caused the change. One way to deal with these methodological challenges is to combine the analysis of documents with interviews and to select multiple key informants per case to minimise biases and ensure triangulation.

Table 5.1 *Potential effects of the accountability mechanisms*

Use and influence of goals and indicators	**Associated concepts**	**Typical stakeholders**	**Intended / positive effects** (declared by proponents or plausible according to literature)	**Unintended / negative effects:** (plausible according to literature)
Instrumental	• **Management** • Decision-making • Policy performance Policy alignment	• Elected politicians • Senior civil servants • Policy experts	• Evidence-informed policy-making • More effectiveness and efficiency • Less clientelism, less corruption	• Rigidity • Gaming • Threshold effects • Costs of reporting
Conceptual	• **Organisational change** • Knowledge, ideas • Networks •	• Staff in local government • Civil society activists / contributors	• Better informed staff, data literacy • Better informed citizenry	• Blame culture • Aversion to learning and experimentation
Political, symbolic and discursive	• **Communication, participation, norm-setting** • Agenda-setting • Legitimacy •	• Citizens • Media • Businesses • Civil society groups	• More citizen participation and trust • Public discourse geared towards sustainability • More transparent political culture	• Public discourse oriented to short-term goal achievement • "Tyranny of light" • Symbolic "greenwashing"

5.3 Methodology and selection of cases

To investigate how the accountability mechanisms play out over time in different contexts, we opted for a longitudinal comparative case analysis using one case from each of four countries (cf. Table 5.2). Such a research design is appropriate for exploring differences and commonalities in practices and different settings and warrants selecting cases on the basis of 'diverse' and 'influential' sampling criteria (Gerring, 2007). Among the countries screened, Colombia is currently the only one with constitutional performance accountability rules for all subnational governments; in our analysis we included the capital (Bogotá) as case study city in the light of its highly influential citizen observatory 'Bogotá Cómo Vamos' (Scrivens & Iasiello, 2010). In Brazil and later Argentina, performance accountability rules were introduced via by-laws at the municipal level. From each country, we selected the city with the longest track record and some academic documentation, namely São Paulo (Cáceres, 2014; Marin, 2016) and Córdoba (Romanutti,

2012b). In Mexico, mayoral performance accountability is not a legal requirement but has been introduced as a voluntary arrangement in two metropolitan areas (Monterrey and Guadalajara) by local civil society organisations (Niemann & Hoppe, 2021). Guadalajara has been subject to prior research (Soto-González, 2020), which justified its selection as a case study. In both Paraguay and Peru, at least one local government has bylaws on *plan de metas* yet these are hardly mentioned in municipal or local media reports, suggesting little active follow-up and limited research opportunities.

Table 5.2 *Prevalence of sub-national performance accountability mechanisms in Latin America*

Country	Government level targeted for performance accountability	Name of the policy (*local name*)	Level of codification, year of origin
Argentina	Mayors in 5 cities (Romanutti & Cáceres, 2020)	Target Plan (*Plan de metas*)	Municipal bylaws, emerging since 2011
Brazil	Mayors in approx. 60 cities (Gaspardo & Ferreira, 2017)	Target Plan / Programme (*Plano / Programa de metas*)	Municipal bylaws, emerging since 2008
Chile	Regional governors (Bravo, 2019)	Programmatic vote (*Voto programático*)	National law (2018); policy not yet applicable to mayors
Colombia	All subnational governments including all governors and mayors (Rey Salamanca, 2015)	Programmatic vote (*Voto programático*), in Bogotá also citizen initiative "Council How Are We doing" (*Concejo ¿cómo vamos?*)	Constitutional article (1994)
Mexico	Mayors in approx. 8 municipalities in metropolitan areas of Monterrey and Guadalajara (Niemann & Hoppe, 2021)	'Mayor, how are we doing?' (*Alcalde, ¿cómo vamos?*) and "Mayor, what have you done?" (*Alcalde, ¿qué has hecho?*)	Voluntary commitments by mayoral candidates, emerging since 1999
Paraguay	Mayor in at least one city (San Lorenzo)	Target Plan (*Plan de metas*)	Municipal by-law introduced in San Lorenzo in 2014; basic reporting only

Peru	Mayor in at least one city (Trujillo)	Target Plan (*Plan de metas*)	Municipal by-law introduced in Trujillo in 2014; no recent reports
Uruguay	Rules for mayors lobbied for by civil society initiatives (Niemann & Hoppe, 2021)	Target Plan (*Plan de metas*)	No approved policy in any city

Our sample of four cases was researched through various data sources (Yin, 2018). In the absence of academic literature, this mainly concerned secondary material such as government websites, NGO reports and blogs, academic theses, media items, YouTube recordings and evaluation studies. In addition, we conducted 16 semi-structured interviews (face to face and via videoconference, lasting 45 to 90 min) between 2014 and 2022 with at least two key informants per case from different stakeholders, such as government representatives and civil society activists. In the case of Colombia, we also interviewed researchers familiar with the national law. The topic guides for interviews and the functions of interviewees are listed in Supplementary materials.

To avoid biases and the emergence of socially desirable answers, the semi-structured interviews with key informants did not probe for the presence of particular benefits. Instead, key informants were interviewed using open-ended questions (about positive and negative effects in the realm of elected politicians and decision-making, local government staff and organisational learning, and citizen trust and participation. To explore how positive and negative effects relate to design factors as well as the socio-political context of each city, we asked key informants about the suitability of mayoral performance mechanisms in different contexts and recommendations for change and improvement. In addition, we analysed the socio-political setup of each city (making use of certain proxy indicators such as country corruption ratings). For national comparisons in other world regions, one may use governance indicators (Glass & Newig, 2019), reliable micro-data are not available at the subnational level for Latin American countries with frequent political upheavals.

For each case, we set up a file containing secondary material and interview transcripts that were translated from the original Spanish and Portuguese into English. The resulting research material was coded and analysed in line with the research framework outlined above. The main characteristics covered were the mechanism's impact on the electoral cycle (pre-election, planning and reporting) and three dimensions of accountability (information, discussions, consequences). Data and emerging patterns were discussed between the researchers to arrive at an indicative rating on the scale high/medium/low on each of the

accountability cube's three dimensions (information, discussions, and consequences) and to arrive at summary statements about each mechanism's strengths and weaknesses. The mechanism's potential effects were researched by listing all declared objectives as found in preambles to municipal by-laws or explicitly mentioned in blogs and policy papers written by advocates. Declared objectives almost exclusively had positive connotations (such as "better informed citizenry," "more efficient public management"). Negative or 'perverse' side-effects (e.g., rigidity, gaming, blame culture) were not mentioned but constitute a possibility according to ample scholarship on performance management. In line with the assessment framework (cf., Table 5.1), we clustered potential effects in terms of the three main uses and influences of performance information (instrumental, conceptual, and political/symbolic/discursive). Results and emerging patterns per case were subsequently compared and condensed to show commonalities and differences between the cities and to help formulate hypotheses for further research.

5.4 Results

To provide essential context information, Table 5.3 presents a summary data of the four case studies. Bogotá, Córdoba, Guadalajara and São Paulo are metropolises with more than 1 million inhabitants, and democratically elected mayors (holding office for 3-4 years) and city councils. In terms of world rankings, Argentina, Brazil, Colombia, and Mexico occupy average places on the UN's Human Development Index and Transparency International's corruption index. According to World Bank assessments, Brazil, Colombia, Mexico, and to a lesser extent Argentina are relatively advanced in terms of results-based public management and monitoring and evaluation capacities (Feinstein & Moreno, 2015).

Table 5.3 *Case study cities and countries*

Case study city	Inhabitants	Mayoral term	Country	Human Development Index (2022)	Corruption Perceptions Index (2022)
Bogotá	8.0 mio	4 years (no immediate re-election)	Colombia	0.76 (rank 91)	39/100 (rank 91)
São Paulo	12.4 mio	4 years	Brazil	0.76 (rank 89)	38/100 (rank 94)
Córdoba	1.6 mio	4 years	Argentina	0.85 (rank 48)	38/100 (rank 94)
Guadalajara	1.4 mio	3 years (with 1 re-election)	Mexico	0.78 (rank 77)	31/126 (rank 126)

5.4.1 Formal characteristics of the performance accountability mechanisms per city

In this section, we present each case study city and summarise our findings in an overview (Table 5.4). For each case, we sketch the country context and how the accountability mechanism affects the electoral-administrative cycle (i.e., election, planning, reporting), followed by a concise description of how the mechanism has been implemented over time. We added an indicative rating of the mechanism in terms of the Accountability Cube's three dimensions.

- Case 1: **Bogotá (Colombia): Evolution of *Programmatic Voting* since 1994**

At a time when Colombia was internationally known for internal strife and crime, Colombia's 1991 constitution ushered in greater levels of decentralisation, a revamped planning system, and various direct democracy mechanisms, including referenda. In 1994, the Programmatic Voting Law came into force, introducing performance accountability to the (sub-national) elections, government plans, and reports. In the election phase, the law requires all candidates for mayor (and governor) to submit their manifestos to a national registry. The National Planning Authority issued guidelines for candidates on how to elaborate such plans (including goals and indicators) but as yet there are no formal content requirements (Muñoz Ocampo & Álvarez Méndez, 2023). After an election, the elected mayor must elaborate a 'development plan' aligning to the manifesto as well as multi-year (provincial, national) master plans and submit this within four months to the municipal council, which has the right to revise, request corrections, and approve the plan. A local and national planning council monitors coherence with existing (master) plans and subsequently of goal achievement (Romanutti, 2012a), for which the mayor must elaborate yearly progress reports.

Amendments to the development plan are possible through the municipal council. Citizen oversight committees are legally mandated to exercise vigilance over all sorts of public management activities (Pogrebinschi, 2023). In recent decades, Bogotá's various mayors have all complied with the Programmatic Voting laws planning and reporting requirements. Importantly, the law foresees the possibility of a recall referendum if, after 12 months of tenure, a mayor is accused of falling to achieve stated goals (Uribe Mendoza, 2016). Recalls are rarely successful (as they require support by 40% of voters) but are frequently attempted. Colombia's capital is home to the highly influential, award-winning civil society initiative *Bogotá Cómo Vamos* (Scrivens and Iasiello, 2010). In addition to compiling and publishing data on sustainability indicators, it organises thematic discussions with mayoral candidates, issues multiple policy recommendations for the development plan, and runs the initiative "Council, How are we doing" that publicly monitors the voting behaviour (including task fulfilment) of municipal councillors. In terms of the three dimensions of accountability (Brandsma and

Schillemans, 2012), the substantial amount of planning and reporting and close scrutiny by civil society and governmental bodies leads to Bogotá's mechanism being rated as 'much' for the information phase and 'intensive' for the discussions phase. The constant threat of a recall or destitution, which came into play when Bogotá's former mayor and later president was temporarily removed from power in 2014, warrants qualifying the sanctions phase as 'many'.

Table 5.4 *History and main features of performance accountability mechanisms in Bogotá, São Paulo, Córdoba and Guadalajara*

	Bogotá (CO): Programmatic Vote	**São Paulo (BR):** Target Plan	**Córdoba (AR):** Target Plan	**Guadalajara (MX):** What Have You Done, Mayor
Year of initiation	1994	2008	2011	2015
Legal framework	Constitution and national law	City-level bylaw	City-level bylaw	Voluntary commitments
Pre-Election **Promises:**				
Content requirements	Alignment with multi-year (provincial) master plans; no detailed content requirements	No formal requirement	No formal requirement	No formal requirement
Submission modality	Registration of all manifestos by national agency	No formal requirements; voluntary public debates	No formal requirements; public debates	NGO collects manifestos and organises public debates
Post-election **Plans:**				
Content requirements	Must align to • multi-year (provincial, national) master plans • manifesto / election promises	Must • align to election promises, Strategic Plan, Budget • contain indicators about sustainable development + promotion of equity, etc	• No reference to election promises • must contain quantified indicators for areas of gov. action (e.g. public works, culture)	• Must by law align to multi-year (provincial, national) master plans • Voluntary commitment to civil soc.: plans for 5 dimensions
Dissemination / deliberation	Submission to and approval by municipal council	Wide dissemination via media and public hearings at district level	Submission to municipal council that organises a public hearing	To public, media
Timing	Within 4 months	Within 90 days	Within 120 days	Not addressed explicitly
Potential adjustments	Via amendment in council	By public dissemination	By informing council and public dissemination	Not addressed explicitly
Post-governing **Reports:**				

Content requirements	Progress	Indicators and goal achievement	Goal achievement	Goal achievement
Submission modality	To municipal council	Public dissemination	To municipal council	To public, media
Timing / periodicity	• Annual progress reports	• Semi-annual indicator reports • Annual progress reports	• Annual progress reports	• Annual progress reports
Accountability Cube:				
Information phase	Much (embedded in nationwide government system)	Medium (quantity of data subject to willingness of politicians)	Medium (quantity of data subject to willingness of politicians)	Little (quantity of data subject to willingness of politicians)
Discussion phase	Intensive (close scrutiny by civil society and gov. bodies)	Intensive: (close scrutiny by large civil society platform)	Intensive (close scrutiny by civil society platform)	Medium (scrutiny by and cooperation with civil society platform)
Consequences/Sanctions	Many (risk of recall)	Medium (risk of municipal censure)	Medium (risk of municipal censure)	Few (no formal sanctions)

- Case 2: **São Paulo (Brazil): Evolution of *Target Plan* since 2008**

After the end of the dictatorship and with the introduction of its 1988 constitution, Brazil became "the world's largest laboratory of democratic innovations" (Pogrebinschi, 2021). Against the backdrop of widespread practices of clientelism and patrimonialism (Gaspardo & Ferreira, 2017), the country obtained an institutional architecture of citizen participation and deliberation and newly empowered municipalities developed instruments such as 'participatory budgeting' that spread internationally (Porto de Oliveira, 2017). In São Paulo, a set of business and civil society leaders founded in 2007 the *Nossa São Paulo* network that, inspired by *Bogotá Cómo Vamos*, started to conduct educational campaigns, publish sustainability indicators, establish work groups about public policies, and monitor the city council actions. The running costs exceeded $1 million annually (Fiabane et al., 2014). Further, the network drafted and successfully lobbied for the 'Target Plan' bylaw approved in 2008. This bylaw does not address the election phase, during which civil society organisations organise debates with candidates willing to participate, but obliges elected mayors to complete a governmental plan within 90 days of assuming power.

Content-wise, the plan must contain the main goals presented in the mayor's electoral campaign, be geared towards "the promotion of environmentally, socially, and economically sustainable development", and have quantified "performance

indicators"; process-wise, the bylaw stipulates: i) wide dissemination in district-level public hearings; ii) the possibility of amendments if changes are required; and iii) semi-annual indicator reports plus yearly progress reports. Non-compliance with these formal rules is punishable by impeachment, yet there are no sanctions for non-achievement of targets. Since the bylaw came into force, all mayors have complied with formal requirements. The first administration (2009-2012) produced a plan with 223 targets, and the next one (2013-2016) built a dedicated website for a plan with 123 targets and links to the budget that was discussed in 35 public hearings, reaching cumulatively more than 10,000 participants (Baliña, 2017). The third administration (2017-2020) received more than 20,000 proposals during the public consultation phase of its plan (Cidade de Sao Paulo, 2022) yet it deactivated the previous website and faced several implementation difficulties due to the Covid-19 pandemic.

Following São Paulo's leadership, similar bylaws with minor adjustments were established in over 60 other Brazilian municipalities (Gaspardo & Ferreira, 2017). Attempts by supporters to turn the Target Plan into a national law were thus far unsuccessful (Programa Cidades Sustentáveis, 2020), just as lawsuits by opponents who had argued that Target Plans interfere with the executive's constitutional independence (Villela & Souza, 2021). Compared to Bogotá, the 'information phase' (Brandsma & Schillemans, 2012) of São Paulo's accountability mechanisms is considered more limited (rating: medium) yet the 'discussion phase' is considered equally intensive. The 'consequences' are rated as 'few' as there are hardly any sanctions – neither legally nor politically – for non-compliance.

- Case 3: **Córdoba (Argentina): Evolution of *Target Plan* since 2011**

In contrast to Brazil, Argentina transitioned to democracy in the 1980s without a substantial constitutional reform. The 1994 amendment increased some elements of citizen participation that have since evolved unevenly across the country, with two-thirds of participatory innovations observed in some large cities (Pogrebinschi, 2021). In Córdoba, Argentina's second largest city, a coalition of universities, business leaders, media firms, and civil society representatives established the network *Nuestra Córdoba* ("Our Córdoba") in 2009. Inspired by its namesake in São Paulo yet institutionally independent, *Nuestra Córdoba* begun collecting and publishing quality-of-life indicators, mobilising citizens, and lobbying for sustainable development policies. Furthermore, the network managed to gain approval in the municipal council of a Target Plan modelled after São Paulo's. Córdoba's 2011 bylaw, too, does not address the election phase, yet Argentina's national laws mandate the diffusion and public debate of election manifestos (Cruz Pérez, 2024).

In Córdoba, the elected mayor must elaborate and submit a Target Plan to the municipal council within 120 days. The municipal council, and not the mayor himself or herself, is also tasked with the plan's diffusion. Contrary to São Paulo, the bylaw in Córdoba does not explicitly reference election promises or to the promotion of 'sustainable development'. Instead, the target plan is required to cover main lines of action (short and long-term) for at least "public works and services, public administration, health, social action, environment, culture, education, tourism, neighbourhood participation and economic development", in each domain with "quantified indicators". The plan can be adjusted by informing the council and the public, and the mayor must submit yearly progress reports. A lack of goal achievement is not sanctioned. However, the lack of reporting constitutes a "serious irregularity".

Since the bylaw entered into force, all mayors formally complied with it. The first administration presented a plan with 513 targets (for 214 objectives in 16 thematic areas). According to civil-society observers, 42% of these targets could not be monitored due to the lack of suitable indicators and baseline values (Romanutti & Cáceres, 2020). The next administration produced a plan with 397 targets for 22 objectives and easier access to information thanks to the use of an open data platform. Public hearings were merely of an informational nature and had no deliberative character yet they still achieved higher levels of public interest than other hearings (Romanutti & Cáceres, 2020). The network *Nuestra Córdoba* computed overall goal achievement levels that were significantly lower than the percentage communicated by the mayor (Baliña, 2017). Following the lead of Córdoba, similar target law bylaws with minor adjustments were established in several other Argentine municipalities (Gaspardo & Ferreira, 2017). In terms of the 'accountability cube' dimensions, Córdoba's mechanism fares the same as São Paulo's, with 'few consequences' yet 'medium-level information' and 'intensive discussions' due to the continuous engagement by the civil society network that had conceived it.

- Case 4: **Guadalajara (Mexico): Evolution of *What Have You Done, Mayor* since 2016**

Although Mexico's constitution had not undergone major changes since 1917, the country initiated decentralisation reforms in the 1980s and transitioned from a de facto one-party regime to multi-party democracy in 2000. In terms of national-subnational coordination, Mexico has a multi-layered system resembling that of Colombia, with local governments crafting 'development programmes' that need to be aligned with higher-level, national plans. In recent decades, Mexico has also seen the emergence of numerous consultative bodies, giving it the most participatory institutions in the region (Pogrebinschi, 2021).

Several citizen-run transparency initiatives concern the country's deteriorating levels of organised crime and impunity. In Guadalajara—the country's third-largest metropolitan area and capital of Jalisco state—a group of business leaders, nonprofit organisations, and universities took inspiration from the example of *Bogotá Cómo Vamos* and founded *Jalisco Cómo Vamos* in 2010. This organisation, primarily a citizen observatory, monitors and disseminates local indicators and public policy proposals through its website and media contacts.

In 2015, *Jalisco Cómo Vamos* started the "What have you done, Mayor" programme as a voluntary arrangement: In the run-up to municipal elections, mayoral candidates were invited to debates, to make explicit manifestos, and to commit to participating in the programme if elected. Once elected, the mayor elaborated—in collaboration with a technical team from *Jalisco Cómo Vamos*—a plan with goals and indicators (currently in five domains "economy, education, environment, security, urban development, and public services"). During a mayor's tenure, *Jalisco Cómo Vamos* tracks goal achievement and disseminates this once or twice per year via a dedicated website and public events. When first run in 2015, the mayors of two municipalities, including Guadalajara City, participated. In subsequent electoral cycles, four other municipalities from the metropolitan area joined. Compared to the other case studies, the 'information' phase of this entirely voluntary accountability mechanism is rated as relatively 'little' and the consequences as 'few'. The discussion component is rated as 'medium' considering the active deliberations between the citizen observatory and mayors.

5.4.2 Positive and negative effects of the mechanisms

In response to the second research question, we present evidence of instrumental, conceptual, and political-discursive use and influence, along with an analysis of strengths and weaknesses of each accountability mechanism. Table 5.5 presents a comparative summary of the main findings for each case. For each illustrative quote, there is a code pointing to the key informant (cf., List of Interviewees).

- Case 1: **Bogotá (Colombia): Effects of *Programmatic Voting***

According to the law's first article, programmatic voting is "the participation mechanism through which citizens [...] impose as a mandate on the elected person in compliance" with their election promises. Key informants agreed that the law generally has the desired effect of contributing to transparency and to programmatic election debates and evidence-informed policies. These benefits are contingent on the law's integration into a nationwide planning system that is equipped with better quality data (on social, environmental and economic indicators) than that of other Latin-American countries. Local stakeholders consider data availability as a key ingredient that allows the political discourse and monitoring of goal achievement to focus on sustainability outcomes (e.g., air quality) rather than about outputs (e.g., number of trees planted). In Bogotá, the active

work of civil society organisations (such as *Bogotá Cómo Vamos*) in producing quality-of-life data, mobilising stakeholders and proposing public policies was credited as contributing to citizen participation and sustainability-oriented policymaking by successive city governments. Some informants were critical of the set-up: "*Bogotá Cómo Vamos* has lost some of its legitimacy. It is an oversight association of business origin, not a citizen initiative" (CO3). Furthermore, relative improvement over time can lead to a decrease in consensus. In the words of a senior municipal employee (CO3): "When in the 1990s Bogotá was in a serious crisis it was easy to agree about acute issues we had to resolve together. As we improve, there begins to be a divergence of opinion on the city's priorities."

The programmatic voting law's harsh sanctions are also viewed as problematic. As the interviewee (CO3) commented:

> Programmatic voting is theoretically nice but doesn't work well in practice. That candidates have to guarantee something is good. If they are very committed to an issue, they make it explicit in the manifesto and later in the development plan. However, if they do not want to take risks, they set general goals that no one can charge them for.

To remedy this situation, there are calls for legal guidelines on the content of manifestos (Muñoz Ocampo & Álvarez Méndez, 2023). Other critical design factors concern the limited terms of office without re-election ("when the mayor arrives, the first six months are spent preparing the development plan"; CO4) and the fact that the burden of accountability is placed exclusively on mayors. As put by informant CO3:

> People are politically uneducated and don't realise that the mayor cannot govern without councillors. The citizens' logic for electing councillors is one and for mayors another. In fact, the elections are on different days! In large cities, part of the council vote is a vote of opinion but the bulk is a clientelist vote. [...]. One would need a political reform guaranteeing that mayors will govern with a majority.

In this context, key informants interviewed saw the civil society project of monitoring the councillors' work as a positive contribution yet also an expensive undertaking that relies on the continuous investment of resources by *Bogotá Cómo Vamos*.

- Case 2: **São Paulo (Brazil): Effects of the *Target Plan***

São Paulo's bylaw does not stipulate objectives, but a recent evaluation compiled by the municipality describes the Target Plan as a "big pact of transparency between the municipal administration and population" designed to "facilitate and foster social control" (Cidade de Sao Paulo, 2022). According to various municipal and civil society informants, the law contributed overall to more transparency

and programmatic political debates (during elections) and goal-oriented policy-making during a mayor's tenure. The bylaw also triggered significant citizen participation (as evident from public hearings and the initiative's name recognition). Key informants stressed that both types of benefits would not have been achieved without the continuous investment of time and energy on behalf of civil society actors (belonging to the *Nossa São Paulo* network) who were also responsible for the bylaw's initial formulation and spread to other cities. Cooperation in a network of cities lends some stability but did not prevent shocks when one administration (2017-2020) deactivated the municipality's Target Plan dashboard. In the words of a local interviewee (BR1): "It was a big step back when they took down these websites. It was seen as an assault on this idea of the Target Programme and government accountability". Later administrations rebuilt dashboards (and additional tools for geo-referenced monitoring), yet the incident showed the mechanism's dependence on politicians' willingness to engage.

In terms of conceptual effects, the network and its bylaw are also credited with fostering multi-stakeholder partnerships: In São Paulo, "one can observe religious community leaders discussing accountability and transparency, leftist militants discussing indicators and targets, and businessmen demanding opportunities for participation" (Fiabane et al., 2014, p. 834). Similarly, the municipal staff interviewed explained that the drafting of a plan – after the election – led to more intra-departmental coordination, policy coherence, and data literacy among government staff. Negative conceptual effects such as organisational competition or blaming were not observed. However, a study of decision-making practices in the municipality found evidence of three types of negative effects typically associated with performance management: "external gaming, myopia and lock-in" (Marin, 2016). In one instance, São Paulo's mayor decided to stick to the target of constructing waste collection points (against better judgement available a year after planning) because he feared the political costs of amending the target. As commented by a municipal employee (BR1): "If you have to change your targets, then you are seen as someone who is not a good mayor." According to some informants, the fact that most targets in the plans of successive administrations concerned tangible municipal outputs can also be interpreted as a negative byproduct of the entire accountability mechanism that puts the focus on short-term goal achievement (expressed in percentages) rather than city-level sustainability outcomes in the long term.

- Case 3: **Córdoba (Argentina): Effects of *Target Plan***

The preamble to Córdoba's bylaw refers to it as an "instrument of planning and citizen communication". According to key informants, the implicit objectives of better-planned municipal actions and better-informed citizens are generally achieved. There is evidence of high levels of citizen engagement (also due to the communication work done by civil society working in a network with like-minded

groups in other Argentine cities). In terms of instrumental and conceptual effects, municipal employees reported that Córdoba's bylaw triggered better intra-departmental coordination and capacities:

> We had to do trainings on methodological aspects because civil servants were not prepared to formulate indicators. At the beginning it was more of an obligation and now those responsible sit down and plan together. [..]. There are many secretaries who really take it as a basis for measurement and update themselves annually, and check whether their goals are being met (AR3).

In Córdoba's case, the choice of goals was initially the subject of many discussions. One debate concerned the scope of the municipality's influence and accountability. In the words of another municipal employee (AR4):

> We are very focused on product indicators and not outcomes [..]. The dilemma comes as to how far you influence. For example, birth or death rates: You don't know which part was affected by you and which part affected by the province or nation.

In addition, there have been repeated debates about who is mandated to formulate policies when the Target Plan is elaborated after an election. Some municipal administrations have publicly rejected proposals made by civil society actors. As one senior municipal official explained (AR5):

> We did have some interference with Nuestra Córdoba where they confused that their goals should be ours . [...] We strongly insisted that the social contract is the election manifesto, and from there, the goals must emerge

At the same time, the official (AR5) acknowledged that "in the manifesto, the choices should be much more explicit. Because manifestos often describe abstract wishes". Civil society representatives concurred by stating that both manifestos and plans often lack specificity ("They make Target Plans for compliance. Modest ones. And then in the course of their mandate they make strategic decisions", AR1) yet that the entire accountability mechanism led, nonetheless, to a more programmatic electoral cycle:

> These are informal social processes of control and the Target Plans have improved the pre-electoral process in the sense that not just anything is promised any longer. (AR1)

Table 5.5 *Summary of effects of mayoral performance accountability mechanisms in Bogotá, São Paulo, Córdoba and Guadalajara*

	Bogotá (CO): Programmatic Vote (since 1994)	**São Paulo (BR):** Target Plan (since 2008)	**Córdoba (AR):** Target Plan (since 2011)	**Guadalajara (MX):** What Have You Done, Mayor (since 2015)
Declared purpose	"participation mechanism to impose compliance" with manifesto	"big pact of transparency between the administration and citizens"	"instrument of planning and citizen communication"	Tool for "accountability and government-citizen communication"
Instrumental effects				
Intended / positive	• Better informed policies • More transparency	• Better informed policies • More transparency	• Better informed policies • More transparency	• Better informed policies • More transparency
Unintended / negative	• High transaction costs of multiple layers of oversight	• Rigidity, loss of flexibility to adjust policies to changing circumstances	• High transaction costs associated with complex reports	
Conceptual effects				
Intended / positive	• Increased awareness among politicians and other stakeholders of need to cooperate for long-term sustainability outcomes	• Increased data literacy among municipal staff • Increased policy coherence among municipal departments • Multi-stakeholder collaboration	• Increased data literacy among municipal staff • Increased policy coherence among municipal departments	Increased data literacy among municipal staff
Unintended / negative				
Political, symbolic, discursive effects				
Intended / positive	Increased citizen engagement	Increased citizen engagement	Increased citizen engagement	Politicians acknowledging importance of accountability
Unintended / negative	Recall threat fosters timid, unspecific election promises	Focus on short-term and tangible outputs	Focus on short-term and tangible outputs	Citizens daunted by complexity of indicator data

Design and context considerations:				
Strengths	• City-level mechanism part of nation-wide planning and data system	• City-level mechanism part of city network and support programme	• City-level mechanism part of multi-city network	• Economical and flexible set-up
Weaknesses	• Lack of guidelines on content of manifestos • Threat of recall unreasonable vis-à-vis veto power of municipal council	• Use of by-law's depends on mayor's disposition • Lack of guidelines on (outcome-oriented) content of plans and reports	• Use of bylaw depends on mayor's disposition • Lack of guidelines on (outcome-oriented) content of plans and reports	• Initiative entirely dependent on candidates' willingness and civil society resources

- Case 4: **Guadalajara (Mexico): Effects of *What Have You Done, Mayor***

According to the project's website, "This accountability and governance tool seeks to evaluate the achievement of government actions and to establish a communication channel between citizens and the government". Increased uptake of the mechanism by candidates and elected mayors (of multiple municipalities), along with substantial citizen interest in the reports produced, are evidence of the mechanism's having political and discursive effects. In this case, in particular, the investment and prior standing of the civil society organisation was a prerequisite for this achievement. As representative of *Jalisco Cómo Vamos* (MX1) explained:

> What we were doing in terms of monitoring was already completely mediatised so to speak. Among academics and authorities, we had a voice that mattered to the media.

The continued willingness by politicians to participate, however, cannot be taken for granted. Reportedly, local mayors stopped cooperating with another, similarly designed initiative in northern Mexico that was overtly critical:

> They pressed too much. The thing is that if you press too hard it doesn't work. It has worked better for us to sit at the table, teach them, and tell them why investing is important (MX1).

In terms of conceptual effects, the mechanism also benefitted municipal capacities and attitudes. In the words of one key informant (MX1): "They still have to work a lot and find it very difficult to differentiate between impact and process indicators" yet "at least we see progress in that they have fewer and better indicators"; municipal staff "now give meaning to being accountable and see it as a need

to respond". According to politicians, collaboration with civil-society observatories has also led to better policies as an instrumental effect of this mechanism. A mayor commented publicly (Jalisco Cómo Vamos, 2022):

> *What you have done, Mayor*, is a programme that challenges and confronts governments. But above all, it helps us be better, set better goals, and achieve better results". "One great idea of *Jalisco Cómo Vamos* was to conduct a survey to find out what [the children] think. [...] Through this study, what we are doing is making all our policies cross-cutting and centred around girls and boys.

In terms of potentially negative effects at the discursive level, another informant (MX2) mentioned the risk of technical (sustainability) information having adverse effects on civic empowerment:

> If we get too much into the technicalities and talk about 'indicators' suddenly it is more complex to reach citizens.

5.5 Discussion

This study set out to investigate the evolution of performance accountability mechanisms in Latin American cities, which have been under-researched yet offer valuable insights for academic debate about effective urban sustainability governance. In the four cities studied, the legal obligations for goal-setting and reporting were associated with an increase in programmatic policies, intra-municipal cooperation, civil society involvement, and citizen participation. An important contingency factor emerging from our research is that in all four case study cities, many benefits were dependent on the investment of civil-society networks (e.g., for data collection, reporting, websites, pro bono consultancies). Approving the legal framework alone may be merely symbolic: In one Brazilian city, the Target Plan bylaw reportedly fell into disuse when the civil society network that had promoted it ceased its operational activity. Overall, the institutional sustainability of the accountability mechanisms was found to be the lowest in Guadalajara and the highest in Bogotá.

Our findings support the assertion that the thriving of citizen observatories requires the existence of favourable legislation as well as the presence of a civil society with a critical mass and stable resources (Díaz Jiménez & Natal, 2014). It is also in line with large-scale evaluations from the sub-continent. One recent study of 3,744 democratic innovations across Latin America found almost 500 targeting the evaluation of policies, which likely indicates dissatisfaction with democracy, feeding both civil society and government attempts to increase legitimacy (Pogrebinschi, 2023). However, many civil society monitoring initiatives are short-lived (Niemann & Hoppe, 2021).

Unintended, negative, or 'perverse' effects of the accountability mechanisms were not observed on a large scale, but some were identified. One example concerned a municipality rigidly sticking to an outdated output target that was perceived as delivering on promises. This indicates a trade-off between accountability and flexibility, which is well-documented in the public management literature (e.g., De Bruijn 2002). In light of this, the observation that Bogotá has relatively more outcome-based goals deserves further attention. After all, outcome indicators are essential for sustainability-oriented monitoring and policymaking (Hák et al., 2016) and less amenable to gaming; on the other hand, publicly reporting on shorter-term output data can improve government performance and citizen trust, which may be vital in underprivileged neighbourhoods and jurisdictions with lower levels of state capacities. This suggests further trade-offs concerning the relative benefits and constraints of outcome and output indicators in different policy sectors and local contexts.

This study enriches international debates (e.g., Valencia et al., 2019) about SDG localisation inasmuch as existing policy recommendations tend to be vague on the matter; one UN-affiliated guidance document (Sustainable Development Solutions Network, 2015, p. 19) states elusively that "as with SDG targets, it is generally preferable […] to track outcomes (or the ends) rather than means. Yet the choice between input and outcome measures must be handled pragmatically." Too many practitioners are still unaware of the potential downsides (in terms of perverse effects) of performance management. For practitioners, an important recommendation is probing for trade-offs concerning accountability and flexibility and dilemmas in the choice of indicators; outcome-based targets foster long-term, holistic policymaking yet output targets align more easily to local government competencies and citizen demands..

Studying 'frontrunners' is a well-established approach (Gerring, 2007) yet has limitations. In this study, selecting the case of Guadalajara implies a survivorship bias because other voluntary arrangements started in other Mexican cities may no longer exist. Furthermore, a city's targets in terms of policy sector and type was also not assessed systematically in this study. Overall, formal characteristics were relatively easier to code than the effects of each accountability mechanism, which emerge over time and are subject to competing perspectives from different stakeholders (Bovens et al., 2008). We contend that our assessment framework provides a useful starting point for further conceptual work (on additional assessment criteria and context indicators) and more empirical studies.

More research is needed, notably on the relative advantages and disadvantages of directing (city-level, urban) performance accountability and communication towards outcomes or outputs in divergent political contexts. After all, it is known that "good measurement systems and well-functioning performance accountability systems only operate in stable environments with a great deal of

standardisation" (Van de Walle & Cornelissen, 2014, p. 453), yet these are not typical characteristics of cities in Latin America (and elsewhere). Further, in terms of citizen communication, it has been argued that to instil trust in institutions, storytelling may be more effective than "scientific evidence" (Pollitt & Bouckaert, 2011). This too requires further research. We recommend studying a larger sample of cases/cities, with attention to government competencies and resources (cf., Fuhr, Hickmann, and Kern 2018), cities' governance settings (Morita et al., 2020), and the role of indicators in different sectors (environment, health, education) since these bring different dynamics in terms of politics and lay understanding (Batley & Mcloughlin, 2015). To facilitate further research, Table 5.6 lists a set of hypotheses derived from this explorative study and the city from our sample that gave rise to the formulation of each hypothesis.

Table 5.6 *Hypotheses for further research*

Factor	Hypotheses	Relevant case(s)
Design features of accountability mechanism	Use of outcome indicators over output indicators in mayoral plans reduces risks of short-term gaming	Bogotá / Colombia
	High threat of sanctions for non-compliance discourages the elaboration of ambitious manifestos	Bogotá / Colombia
	Reporting on high content indicator data is unsuitable to engage citizens	Córdoba, Guadalajara
Local context and process	The engagement of strong local civil society organisations facilitates the effective implementation of mayoral accountability mechanisms	multiple
	Civil society engagement in performance monitoring by a small set of organisations increases the risk of them being perceived as partisan	Bogotá, Guadalajara
Country-level context	The availability of reliable (national) datasets is a precondition to effectively use outcome-level goals at city level	Bogotá / Colombia
	Broad legal competencies (covering health, education, environment, etc) as well as longer mayoral terms (>4 years) facilitate outcome-oriented performance accountability	multiple
	Strong performance accountability has relatively more benefits in countries with low government effectiveness	multiple

5.6 Conclusion

This study explored the characteristics and effects of novel performance accountability mechanisms that have evolved among pioneering local governments in Colombia, Brazil, Argentina, and Mexico. In response to our first, descriptive research question on the mechanisms' characteristics, we can conclude that the comparative analysis of Bogotá, São Paulo, Córdoba and Guadalajara showed three main types of legal approaches: a constitutional obligation in Colombia,

city-level bylaws in Brazil and Argentina and, voluntary submission to external scrutiny by Mexican mayors.

The mechanisms differ greatly in terms of scope. The Colombian system and Mexican project cover the entire electoral-administrative cycle, from election promises by candidates to a mayor's reports on goal achievement. By contrast, the Brazilian and Argentine bylaws leave out the election phase. The latter approach preempts one difficulty inherent in the Colombian model: the mismatch between candidates making promises as individuals (or parties) and elected mayors needing to govern with competitive, multiparty councils. Another finding—which is not imposed by legal rules—is that many goals and the associated reporting system in Bogotá concern long-term sustainability outcomes (such as quality-of-life). In the other cities, relatively more goals and reports are about the delivery of tangible, short-term outputs such as public works.

Compared with the description of formal features, analysing the effects of an accountability mechanism is highly complex. As such, findings from this comparative case study do not provide comprehensive evidence that would allow us to answer our second, evaluative research question conclusively. However, several noteworthy patterns emerged. In all four cases, the mechanisms showed effectiveness in reaching the attention of citizens and politicians and informing specific (sustainability) policies. Furthermore, there are indications of the mechanisms contributing to transparency and a more programmatic, policy-oriented election cycle with a more participatory process of political deliberation. Positive effects were also reported—at least in the first years after the mechanism's introduction—in the realm of data and indicator literacy, as well as in institutional cooperation among local governments and civil society actors.

What is the current verdict on the dichotomous question in this paper's title: Should mayors be made accountable for their electoral promises? In answering this question, there appear to be trade-offs between accountability mechanisms bringing benefits in government performance and citizen participation, yet costs in terms of inflexibility and 'perverse effects'. Forced to choose an answer, this research suggests a tentative 'yes' for cities with low-performing municipal administrations—some degree of accountability is associated with more programmatic, less clientelist policymaking. In high-performing institutional contexts, rigid performance accountability seems less appropriate. In both types of context, effective sustainability governance requires (i) working towards long-term objectives and outcomes rather than short-term election promises and (ii) exploring alternatives to the currently dominant model of voluntary goal-setting.

6

Synthesis, discussion and conclusions

In today's political context, where calls for 'evidence-based', 'data-driven' and 'sustainable' public management and accountability are commonplace, sustainability reporting is on the rise. In the geographical context of cities, it can take various forms. As explained in the introduction (Chapter 1), real-life examples include formal sustainability reports that some municipalities issue annually as well as less formalised reviews published by nongovernmental 'urban observatories'. The common denominator of these approaches is the disclosure and public dissemination of sustainability-relevant urban trends, political commitments, and actions, with the – more or less explicit – goal of promoting evidence-informed policymaking, accountability, and, citizen engagement. As operationally defined in Chapter 1, city-oriented sustainability reporting refers to periodic, public dissemination of multi-domain information on sustainability performance at the city level, produced by or in collaboration with local authorities or civic actors. However, as the empirical chapters demonstrate, this definition encompasses a broad spectrum of "information-based governance": ranging from collaborative *trend monitoring* (the "State of the City" focus of Chapters 2 and 3) to adversarial *performance accountability* (the "State of the Government" focus of Chapters 4 and 5).Previous studies have demonstrated that sustainability reporting and related information-based urban governance initiatives such as community indicator projects, can yield certain desirable benefits. One can conceptually and empirically distinguish three clusters: i) 'instrumental use and influence', such as effects of reporting on the decision-making processes of policymakers, ii) 'conceptual use and influence', such as organisational learning by government staff and other stakeholders, and iii) 'political-symbolic use and influence' such as changes in the discourse and awareness among citizens, media outlets, and political parties that shape opinions and legitimise actions or inaction. Figure 6.1 summarises these three type of uses and influence, along with their most common stakeholders.

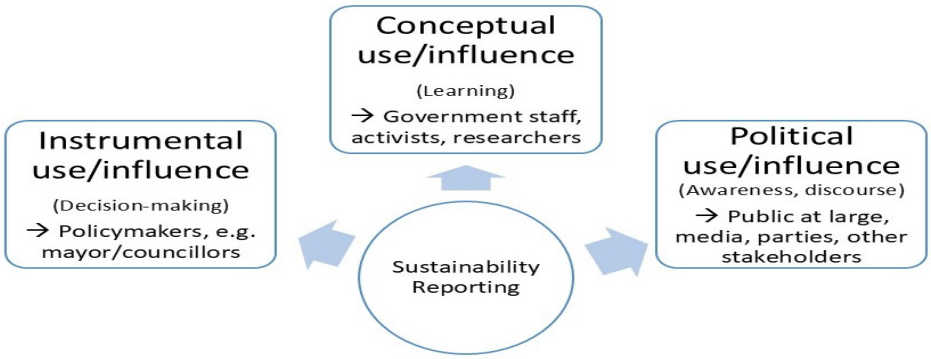


Figure 6.1 *Diverging effects and target audiences of sustainability reporting*

Notably, sustainability reporting is often presented as capable of simultaneously fostering all these positive attributes, regardless of place and time. Are these claims about sustainability reporting as a 'jack-of-all-trades' tool warranted? Are the benefits sustainable over time, and what are potential downsides? What do different forms achieve in different socio-political contexts? These questions have been hitherto under-researched, yet they have high practical relevance, since reporting is a cornerstone of modern governance and promoted by numerous international frameworks including the Sustainable Development Goals (SDGs).

In response to calls for further research, particularly longitudinal case studies, this dissertation investigated the factors leading to the types of success (and failure) of reporting initiatives in diverse contexts. As detailed in the introduction (Section 1.3.2), Europe and Latin America were strategically chosen to enhance the theoretical understanding of city-oriented sustainability reporting and to address significant empirical gaps in some under-researched regions. In each world region, one study assessed initiatives in terms of city-wide indicators, while another focused on the link between reporting and local government performance. Together, these studies illustrate a spectrum of information-based governance ranging from voluntary disclosure to mandatory accountability. The following research question guided this inquiry:

> **How can sustainability reporting initiatives benefit decision-making, learning, and public awareness in the divergent contexts of European and Latin American cities?**

This chapter presents answers to this overall research question about design, contextual factors, and effects. To achieve this, Section 6.1 first provides answers to the sub-questions that guided the four empirical studies presented in Chapters 2-5. These findings are then used in Section 6.2 to address the main research question by identifying 'design strategies' and theoretical mechanisms that help explain the success of reporting initiatives for divergent purposes and contexts. Section 6.3 discusses the main contributions and limitations of this dissertation and offers suggestions for further research. By reflecting on the broader theoretical and practical contributions of the dissertation, a refined framework for evaluating city-oriented sustainability reporting initiatives is presented with the aim of helping future studies. Section 6.4 then presents a summary of practical policy implications derived from this dissertation.

While the introduction (Chapter 1) situated this research within the broader context of existing literature (cf. the research funnel metaphor in Section 1.3.2), this chapter aims to create an 'inverted funnel' (cf. bottom part of Figure 1.8) by expanding the discussion again and reviewing broader implications.

6.1 Answers to the research sub-questions

The four empirical studies of this dissertation share a methodological 'red thread' – as described in introduction (cf. Sections 1.3.4.-1.3.6), they all rely on longitudinal, mixed methods designs with a similar conceptual framework[7]. The subsequent sections are dedicated to synthesizing the findings from these individual studies[8]. In line with the overall framework (cf. Section 1.3.5), this synthesis is based on three factors: (i) contextual considerations, (ii) the design factors of "internal organisation" and "content-related activities," and (iii) evolving effects in terms of "uses and influence" and organisational continuity.

- Sub-question 1: **How have sustainability reporting practices evolved in pioneering European city governments, and what are their effects?**

The research on selected European cities (Amsterdam, Basel, Dublin, Freiburg, Nuremberg, Zurich; described in detail in Chapter 2) shows that sustainability reporting is a growing trend, generally initiated voluntarily.

This voluntary emergence allowed for diverse experimentation with scope, frequency, and design. City authorities crafted reporting systems using a mix of fixed and rotating indicators, aligned with administrative and electoral cycles, and embedded them within strategic planning and budgeting processes. This diversity was shaped by each city's specific governance model and policy objectives, with some, like Freiburg, using reporting primarily as a tool for strategic decision-making by the municipal council, while others, like Amsterdam, adopted more fragmented approaches targeting different audiences through multiple formats.

In terms of context, the studied cities shared certain favourable characteristics: stable institutional settings, relatively high trust in government, and experienced civil servants with capacity for long-term coordination. At the same time, differences in mayoral term lengths and local political mandates influenced the rhythm and continuity of reporting efforts. Although all cities shared a common commitment to sustainability, the motivations behind reporting varied—from internal management to civic engagement.

Most initiatives initially benefited from novelty effects. They gained political attention, media coverage, and stimulated interdepartmental cooperation.

[7] Note on Terminology: This dissertation spans several years of research. The terms 'Monitoring' (used in Chapter 4, 2017) and 'Accountability Mechanisms' (used in Chapter 5, 2025) reflect the evolution of the field from *technical measurement* to *political governance*. In this final synthesis, 'Sustainability Reporting' is used as the umbrella term encompassing both.

[8] In order to gain insights into social mechanisms that transcend individual cases (cf. Section 1.3.2 on 'middle range theories'), it is necessary to make certain deductions and inferences. Concurrently, methodological rigour requires remaining faithful to the empirical data from the heterogeneous case studies and acknowledging the limits to generalisations.

However, these positive effects often proved difficult to sustain. Reporting fatigue set in as the novelty wore off, and both public interest and media visibility declined. In some cities, stakeholder involvement dropped significantly in later stages, raising concerns about long-term relevance and use.

Regarding effects, the reporting initiatives showed a range of outcomes. Instrumental uses were observed in internal management and planning processes. Conceptual effects included increased awareness and interdepartmental learning. In some cases, political and symbolic uses of reporting were evident, with local politicians highlighting selected indicators to signal achievements or justify policy directions. Yet, the durability of these effects was mixed, and their magnitude declined over time. Moreover, the tension between reporting for managerial control and reporting for public accountability became apparent. Cities struggled to design reports that were both technically rigorous and engaging for wider audiences. This trade-off between comprehensiveness and accessibility remains unresolved.

Since the selected cases represent frontrunners with relatively successful and enduring reporting systems, the findings may be subject to survivorship bias and cannot be directly extrapolated to less mature or short-lived initiatives. Table 6.1 provides a comparative overview of these cases, summarising key features of the reporting systems, their institutional context, and the observed effects.

Table 6.1 *Summary findings from research about sustainability reporting initiatives in six European cities*

Factor	Main findings from sample of 6 European cities
Context	• Stable political context; relatively slow urban change; generally high-quality data availability. • Mostly consensual democracies with established bureaucracies and high institutional trust. • Differences across countries in local government mandates, length of mayoral terms, and fiscal autonomy. • Stable context enabled long-term planning and coordination across municipal departments.
Internal organisation	• Reporting developed and maintained voluntarily by city governments. • Often inspired by international or national frameworks (e.g., GRI, Local Agenda 21), even when implementation was locally tailored. • Some cooperation & knowledge sharing between cities (e.g., Zurich + Basel)
Content-related activities	• Annual or multi-annual reports, often with thematic rotations or fixed indicator sets. • Limited use of standardised frameworks; cities adapted global indicators to local conditions. • Focus on outcome indicators and reporting on municipal actions. • In some cases, integration with long-term planning and budgeting (e.g., Freiburg, Nuremberg).

	• Tension between technical depth and public accessibility: reports were often detailed but less engaging for wider audiences.
Organisational continuity	• Multi-year continuity of reporting efforts across all cases, reflecting sample selection bias toward durable initiatives. • However, signs of 'reporting fatigue' emerged over time, including reduced stakeholder interest, declining media coverage, and limited novelty in later reporting cycles
Evolving effects	• Evidence of instrumental (internal planning), conceptual (learning and awareness), and political (symbolic or strategic) effects in varying degrees. • Effects were typically strongest in early stages of reporting initiatives and tended to diminish over time without innovation or renewed engagement strategies

- Sub-question 2: **How have sustainability reporting practices by community indicator initiatives evolved in selected Latin American cities, and what are their effects?**

The research on Latin American cities – described in Chapter 3 – showed that city-oriented sustainability reporting systems experienced rapid growth in the previous decade, reaching 70 cities in ten Latin American countries compared to half that number five years earlier. This growth was driven not only by policy diffusion efforts among civil society groups but also by inter-governmental donors, particularly the Inter-American Development Bank (IADB). The 10 Latin American countries that are home to the 49 reporting initiatives studied differ in terms of population size and socio-economic development. Compared to Europe, however, the political context is more majoritarian and adversarial, with several countries suffering from high levels of violence, corruption, and weak state capacities—including limited data availability.

As summarised in Table 6.2, most of the successful reporting initiatives in Latin America were driven by local coalitions of civil society organisations, universities, media, and philanthropic foundations. These alliances—either close-knit or network-based—typically excluded government actors, which distinguishes them from similar initiatives in other world regions. These organisational forms varied in their structure and durability: close-knit models tended to feature clearer mandates and branding, while networks were more adaptable but potentially less institutionalised. In countries like Colombia, recognisable umbrella networks such as *Cómo Vamos* enabled both resilience and visibility.

Table 6.2 *Summary findings from research about sustainability reporting initiatives in 49 Latin American cities*

Factor	Main findings from sample of 49 Latin American initiatives in 10 countries
Context	• Unstable political context; fast urban growth and high inequalities, coupled with poor data availability.

	• Mostly majoritarian and clientelist politics; civil society and transparency initiatives often face adversarial relations with local authorities. • Cross-country variation in local government mandates, mayoral term lengths, and the institutionalisation of data systems.
Internal organisation	• Reporting typically led by civil society coalitions involving NGOs, universities, media outlets, and philanthropic foundations. • Organisational forms range from close-knit institutional setups to looser city networks. • Recognisable branded initiatives in some countries (e.g. Cómo Vamos in Colombia) helped sustain visibility and credibility.
Content-related activities	• Annual reports and online dashboards • Limited use of standardised indicator frameworks • Focus on multi-year outcome indicators and local government actions; collection of subjective well-being data via population surveys • Training activities on 'data literacy' for government and civil society • In some cities, attempt to link to long-term planning / budget • Use of annual reports and online dashboards. • Limited reliance on standardised indicator frameworks; strong emphasis on outcome indicators and perceived performance. • Extensive use of citizen perception surveys for public engagement and media visibility. • Use of proxy indicators and FOIA-based data acquisition to address unreliable or inaccessible official data. • Training activities for public servants, journalists, and civil society on data literacy and sustainability indicators. • In some cities, partial links to long-term planning or budgeting processes.
Organisational continuity	• Multi-year continuity in approximately 55% of initiatives after seven years. • Discontinuation more frequent in very high- and very low-income countries; greater continuity in upper-middle-income contexts with stronger civil society infrastructure.
Evolving effects	• Evidence of instrumental, conceptual, and political influence across initiatives. • Effects were strongest in early years and tended to diminish over time due to 'reporting fatigue', reduced stakeholder engagement, and funding instability. • Notable cases achieved media visibility, policy influence, and agenda-setting power, especially in large cities such as Bogotá and São Paulo.

Several initiatives (notably in Brazil) faced political pressure, particularly when their findings or recommendations conflicted with the interests of local authorities. In some cities, this reduced government cooperation or even led to efforts to

discredit the initiatives. Political interference remains a significant challenge, especially in countries with volatile political environments. Continuity was more likely in politically stable countries with stronger civil society infrastructure and medium income levels, while initiatives in highly unequal or extremely low-income settings were more prone to discontinuation.

Faced with unreliable or absent official data, initiatives often relied on proxy indicators or estimates. In response, some civil society organisations made use of freedom of information requests to improve data access. Another important strategy was to provide stakeholders—including public servants, journalists, and community leaders—with training on data literacy to build promote informed participation.

In terms of content, a prominent strategy for many initiatives in Latin America was the use of citizen perception surveys. Although expensive to implement, these surveys proved effective in garnering media attention and engaging local populations. By assessing public opinion on various sustainability-related topics and disseminating results with media partners, these initiatives raised awareness and held local governments accountable for issues such as child mortality in underserved neighbourhoods.

Several reporting initiatives contributed to organisational learning within civil society organisations and, in some cases, local governments. These conceptual effects, however, tended to be most pronounced in the early years of an initiative's operations and diminished as the initiative matured. Many Latin American initiatives experienced 'reporting fatigue,' especially when the perceived impact of reports waned or funding was insecure. This was particularly evident in countries like Chile, where disengagement followed an initial burst of enthusiasm. Financial instability, often tied to reliance on external philanthropic grants, posed a recurring threat to long-term viability.

Despite these challenges, several initiatives achieved high name recognition and tangible policy influence. This was notable in megacities such as Bogotá and São Paulo, where sustained advocacy contributed to changes in urban planning policy. In some cases, reporting led to the adoption of bylaws, while in others, it helped place issues such as air quality and sustainable mobility on the political agenda. These effects can be understood as contributing to social accountability and organisational learning—both of which underpin the potential of community-driven reporting to strengthen democratic governance.

- Sub-question 3: **What are the strengths and limitations of reporting initiatives addressing the sustainability policies of European local governments via voluntary benchmarking?**

As described in Chapter 4, only a few sustainability reporting initiatives in Europe have focused on benchmarking the implementation of sustainability policies by local governments. Many tools were initially developed for this purpose, but most failed to attract sufficient participation and were eventually abandoned. One notable exception is the Dutch "Local Sustainability Meter" (LSM), which achieved an unusually high degree of staying power.

As summarised in Table 6.3, the LSM's design strategy relied on voluntary self-assessments by municipalities, resulting in benchmarking through publicly visible rankings and awards. Municipalities would rise or fall in the rankings based on how many of approximately 100 prescribed sustainability policies they self-reported having implemented. Launched in 1999, the LSM achieved peak participation from around 90% of all Dutch municipalities in some of its multi-year editions. This case provided valuable insights into how a policy-focused reporting initiative can function over the long term.

Table 6.3 *Summary findings from research about voluntary benchmarking in Dutch cities*

Factor	Main findings from evaluation of Dutch "LSM"
Context	• Stable political context; relatively slow urban change • Netherlands: consensual democracy with high professional discretion for municipal staff and a culture of transparency and self-criticism. • Identical mandates and election cycles across municipalities enabled meaningful benchmarking. • Institutional environment fostered voluntary participation, even by low-performing municipalities
Internal organisation	• Reporting designed and managed by an NGO with arm's-length public funding (1999–2014). • Government support was financial but not directive, enhancing perceived neutrality and legitimacy.
Content-related activities	• Voluntary self-assessment of implementation of ~100 municipal sustainability policies. • Annual and multi-annual rankings, awards for top performers, and rotating thematic focuses. • Emphasis on policy adoption rather than outcome indicators. • Public dissemination through media, events, and municipal reports.
Organisational continuity	• High early uptake: around 90% participation during peak periods. • Declining engagement over time due to reporting fatigue, growing administrative burden, and reduced media interest.

	• Funding and political attention diminished in later years.
Evolving effects	• Instrumental effects: adoption of new policies to improve rank. • Conceptual effects: internal learning, agenda-setting, and inter-municipal collaboration. • Symbolic/political effects via public rankings. • Risk of "greenwashing" due to focus on reportable actions rather than actual sustainability outcomes.

The Dutch context was conducive to the LSM's adoption: municipalities operate under uniform legal mandates, and civil servants enjoy a high degree of professional discretion. The political culture also values transparency and self-reflection. This enabled effective benchmarking and explained the rare phenomenon of low-scoring municipalities voluntarily disclosing their performance.

The LSM encouraged municipalities to review and update their policy portfolios, compare themselves to peers, and identify areas for improvement. The public rankings and accompanying awards ceremonies generated media attention and civic interest, especially during the early years. The tool also encouraged collaboration and learning among municipalities, including the formation of inter-municipal networks on topics such as energy efficiency and waste management.

In terms of influence, the LSM produced both instrumental effects (e.g., policy adoption) and conceptual effects (e.g., increased awareness among municipal staff). However, its design also had limitations. The LSM measured the presence of policies rather than their effectiveness or outcomes. This focus increased the risk of 'greenwashing', where municipalities might emphasise visible, easily reportable actions over substantive environmental performance.

As sustainability reporting became more institutionalised and expectations grew, municipalities faced increased administrative burdens to meet the reporting requirements. Over time, the novelty of the rankings faded, public engagement declined, and some municipalities reduced their involvement. Participant fatigue and rising data demands led to a reduction in support from the Dutch government, which had provided long-standing financial backing. The LSM's success during its early years was partly due to a lack of competing tools and the specific institutional characteristics of Dutch local governance, making it a unique case within the broader European context.

- Sub-question 4: **What are the strengths and limitations of reporting initiatives addressing the sustainability policies of selected Latin American local governments via accountability obligations?**

As described in Chapter 5, several Latin American countries experimented with performance-oriented reporting mechanisms amid conditions of weak state capacity, low public trust, and fragmented governance. Colombia is a frontrunner, where since the 1990s, all mayors have been constitutionally required to align their governing plans with election promises and report on progress during their term. Similar frameworks, often influenced by civil society advocacy, have been adopted in more than 60 cities across the region.

Results of the longitudinal comparison of Bogotá (Colombia), Córdoba (Argentina), Guadalajara (Mexico), and São Paulo (Brazil) are summarised in Table 6.4. Many systems required mayors or municipal departments to produce regular reports detailing goal achievement on the basis of tailor-made formats or legally defined indicators. These reports were subject to public scrutiny by civil society watchdogs or participatory institutions. In some cities, written reports were complemented by public forums or hearings that enabled citizen feedback and strengthened civic participation.

Table 6.4 *Summary findings from research about accountability-focused reporting in Latin American cities*

Factor	Main findings from sample of 6 Latin American initiatives
Context	• Unstable political environments with adversarial or clientelist dynamics. • In several countries, weak state capacity and high corruption levels. • Variability in local mandates, election cycles, and civic oversight institutions across countries. • Civil society filled oversight gaps where formal institutions lacked capacity
Internal organisation	• Reporting conducted by local governments on progress toward election promises and annual targets. • Degree of obligation varied: some mechanisms were constitutionally mandated (e.g., Colombia), others driven by civil society pressure. • Civic watchdogs or participatory bodies scrutinised content and demanded greater accountability.
Content-related activities	• Reporting formats from legally binding plans to locally negotiated goals. • Indicators often focused on short-term outputs; in some cities (e.g., Bogotá), complemented by longer-term sustainability outcomes. • Public engagement mechanisms included hearings, forums, and media coverage to increase transparency.
Organisational continuity	• Formal compliance with reporting requirements was common, but quality and depth of reports varied. • Tensions between elected officials and civil society were frequent, especially in cities with vocal watchdog institutions.

	• Sustainability of these mechanisms depended on institutionalised civic oversight and administrative continuity
Evolving - effects	• Instrumental and conceptual effects included better planning, learning, and civic engagement. • Political effects varied, sometimes enabling reputational incentives or backlash. • Risks included "gaming" of indicators, selection of low-effort targets, and rigid adherence to outdated plans. • Overall effectiveness hinged on the balance between external pressure, data quality, and adaptive use of reporting tools

By linking political commitments to measurable targets, these initiatives encouraged greater transparency and made it easier for citizens and civil society to assess whether local governments delivered on their promises. Several mechanisms promoted the integration of sustainability objectives into local planning documents, enhancing programmatic governance and helping shift focus away from clientelist or short-term policymaking. In their early phases, these systems often supported capacity-building in local administrations, especially in data collection, indicator use, and results monitoring.

However, as the reporting systems matured, their limitations became increasingly visible. Monitoring requirements created administrative burdens, particularly in cities with weak data systems or high staff turnover. In many cases, the availability and reliability of sustainability-related data posed significant challenges. When accountability mechanisms focused heavily on measurable outputs, they risked incentivizing narrow compliance or the selection of vague, easily achievable targets—a dynamic akin to 'greenwashing'. Furthermore, strict adherence to pre-set targets could reduce adaptive flexibility, discouraging course corrections even when conditions changed. This tension reflects what Chapter 5 characterises as the trade-off between productive and perverse accountability: systems may foster learning and reform, but can also create rigidity, symbolic compliance, or superficial success narratives.

The effectiveness of these systems, therefore, depends not only on legal mandates but also on the institutional strength of civil society, the quality of data infrastructures, and the political willingness of municipal leaders to engage in open and adaptive accountability practices.

6.2 City-oriented sustainability reporting: Design choices and ecological fit

The key theoretical interest centres on the question of whether and how sustainability reporting can become and remain useful for cities in divergent contexts. As mentioned before, 'useful' refers to various positive effects, while negative or 'perverse' effects and costs are minimised. Indirectly, the organisational continuity and 'staying power' of a reporting initiative itself can be regarded as a precondition for sustained success. The scoping exercises conducted as part of the four empirical studies showed that in European and Latin American cities, most sustainability reporting initiatives in both world regions were initiated as voluntary commitments by local governments or civil society organisations. Many were discontinued after a few years, presumably sooner than their instigators had initially hoped for. Some, however, achieved staying power and led, to various degrees, to effects that can be considered positive, as the empirical studies have shown. Examples of positive effects include instrumental use and influence (when sustainability reporting contributed to evidence-informed decision-making), conceptual use and influence (via contributions to 'enlightenment' and learning), and political use and influence (such as the promotion of awareness, legitimacy and multi-stakeholder support for urban sustainability).

The empirical studies presented in Chapters 2–5 and summarised in Section 6.1 thus contain partial answers to the overall research question. Most reporting initiatives can be regarded as real-life demonstrations of 'how city-oriented sustainability reporting can be successful'. After all, the various reporting initiatives that were researched and presented–ranging from annual reports by European municipalities to indicator frameworks run by Latin American urban observatories –each produced some positive effects with a limited set of costs and risks. However, such benefits did not persist without adaptation and strategic renewal.

As evident from the four studies, there are several differences among reporting initiatives in European and Latin American cities (cf. especially 'internal organisation' in Tables 6.1-6.4). This precludes a systematic analysis of how similarly constructed reporting initiatives play out in different contexts. Nonetheless, the various characteristics of reporting initiatives can be linked to patterns of effects in relation to contextual variables. In the overall evaluation framework presented in the introduction (cf. Figure 1.7), an initiative's 'design' constitutes the linking pin between the external context (at country and city level) and various types of instrumental, conceptual and political effects. Conceptually and methodologically, it is therefore appropriate to answer the overall research question with an analysis of major 'design strategies'. As also visualised in the evaluation framework (Figure 1.7), one can conceptually distinguish two main clusters of **design choices** that each reporting initiative needs to make:

i. Internal organisational and governance arrangements
ii. Content and dissemination activities of reports.

The four empirical studies produced insights about design strategies in each cluster. Importantly, some of these strategies appear to be 'timeless' and generalizable whereas others are contingent on the particular purpose and context of the reporting initiative. In response to these differences, the synthesis of findings observed across Europe and Latin America proceeds by distinguishing:

i. Widely applicable strategies;
ii. Purpose-dependent strategies (usually involving trade-offs);
iii. Context-specific insights and contingencies.

Based on this distinction, various **design strategies** were identified that are summarised in Table 6.5 and further described in Sections 6.2.1 and 6.2.2. However, these strategies do not operate in isolation. To counter the risk of oversimplification, the theoretical synthesis regarding the configuration and ecological fit of these strategies is discussed in Section 6.2.3.

Table 6.5 *Widely applicable, purpose-dependent and context-dependent strategies*

Design factor	Assessment of strategies
Internal organisation and governance	
Stakeholders of the reporting initiative	Widely applicable: • Cooperation with universities and other reputed (media) organisations Purpose-dependent: • Governmental participation fosters learning yet impairs impartiality • Voluntary participation increases ownership yet risks discontinuity Context-specific: • Potential for cooperation with local government depends on politisation, corruption, capacities, and need for impartiality • Availability of civil society actors / NGO funding and media independence differs by country • Presence of competing "transparency and accountability" tools influences viability
Benchmarking and networking with other cities	Widely applicable: • Networking / branding with other cities helps name recognition • Benchmarking can provide insights & serve as the essence of rankings Purpose-dependent: • Benchmarking requires standardised ways of working which may impair bottom-up, tailor-made approaches Context- specific:

	• Benchmarking requires participation by other, comparable cities (e.g., in terms of sustainability challenges and government competencies)
Monitoring and evaluation	Widely applicable: • Monitoring of media uptake; periodic self-evaluation with stakeholders
Content and dissemination strategies	
Choice of indicators and data sources	Widely applicable: • Standardised indicator frameworks can save resources and enhance legitimacy • Balance between too few (simplistic) and too many (complex) indicators • Use of open data laws for data acquisition • Citizen wellbeing surveys have political salience and accessibility Purpose-dependent: • Instrumental use benefits from alignment with local government competencies and cycles • Outcome indicators reflect long-term trends but may misalign with local government competencies • Output indicators align with local actions but risk short-termism • Process indicators foster innovation but may emphasise effort over results Context-specific: • Government competencies differ per country, with local governments having limited influence in some domains (e.g., climate change) • The availability of reliable data differs by country • Understanding of sustainability by general public differs by region
Accountability for government action and goal achievement	Widely applicable: • Performance accountability can drive performance and enhance trust Purpose-dependent: • High accountability pressure risks producing rigidity and strategic behaviour (gaming) and is difficult to align with multi-stakeholder work Context-specific: • In politicised environments, meaningful commitments may be avoided; civic oversight strength varies
Focus, periodicity, timing of reports	Widely applicable: • Combining trend monitoring with thematic foci enhances interest • Timing with electoral cycles and events increases salience Context-specific: • High-frequency reporting more effective in fast-changing urban contexts (e.g., Latin America)
Outreach and dissemination	Widely applicable: • Round-tables and inclusion of political parties, journalists, etc. • Rotating focus themes maintain engagement

	Context-specific: • Effective channels differ by country (e.g., digital media, literacy, public meeting traditions)

6.2.1 Internal organisation and governance of initiatives

In the cluster of 'internal organisation', one widely applicable strategy is cooperation with multiple stakeholders, especially those perceived as reputable and legitimate (e.g., universities). However, the *depth* of this cooperation must be calibrated to the local level of trust. A key context-specific variable is the local availability of non-governmental stakeholders willing to cooperate, such as universities, NGOs, and foundations, as well as the degree of media independence. While municipal participation can promote organisational learning between governmental and non-governmental actors, it may also compromise impartiality (a critical asset in polarised environments). This presents a trade-off and thus represents a purpose-dependent strategy, as summarised in Table 6.5. The relative merit of this strategy also depends on contextual factors, including a country's degree of politicisation, corruption levels, and the capacities of governmental and non-governmental institutions. Further contextual contingencies of importance include data availability (relatively higher in Europe and lower in Latin America) and the presence of other competing transparency and accountability tools that may draw the attention of city governments and the public.

Another key design choice concerns the mode of institutional participation. Compulsory performance accountability, as seen in several Latin American cities (Chapter 5), can increase the instrumental use of indicators but may lead to low commitment and the 'gaming' of targets. As noted in Table 6.5, strong accountability pressure can, in some settings, lead to strategic behaviour rather than genuine progress. Voluntary participation, by contrast, is more likely to foster organisational learning and can encourage critical self-assessments (cf. Chapter 4 on experiences with the Dutch LSM). However, the effectiveness of a voluntary approach is context-specific, requiring a political-administrative culture that supports openness and critical self-reflection. A related risk of voluntary participation is discontinuity, especially when novelty fades or when reported performance is unfavourable.

Cooperation between cities is a further widely applicable strategy—for example, in promoting name recognition and benchmarking, as demonstrated in Colombia (Chapter 3) and the Netherlands (Chapter 4). However, it also requires the active participation of comparable cities with similar characteristics. As highlighted in Table 6.5, benchmarking is often dependent on standardised procedures, which may limit bottom-up, tailor-made approaches. Furthermore, transnational

networks formed in Latin America and Europe (Chapters 3 and 5) were ultimately unsuccessful, likely due to a lack of common ground among participants.

In this cluster, another widely applicable strategy was found to be the monitoring of effects by reporting initiatives—such as tracking the reach of communication products among local stakeholders—and undertaking periodic self-evaluations. These practices support internal learning, credibility, and adaptation over time, and can help maintain engagement in the face of changing political or institutional circumstances.

6.2.2 Content and dissemination strategies of reports

The second cluster of design strategies concerns the manifest content of sustainability reporting and the modes of reporting—namely, what the reports cover, how frequently they are issued, and how they are disseminated. Overall, the application of standardised reporting frameworks (e.g., derived from the GRI, SDGs, or ISO) was found to be widely applicable across various contexts. These frameworks help save resources, enhance legitimacy, and enable comparability. However, they also carry the risk that certain indicators may not be fully relevant to local conditions. As summarised in Table 6.5, a robust strategy combines the use of standardised indicators where appropriate with the addition of tailor-made indicators where necessary.

Regarding the number of indicators—and thus the complexity of reporting—a middle ground proves generally effective. This avoids the extremes of too few indicators, which may oversimplify, or too many, which can overwhelm or obscure key findings. For data acquisition, leveraging open data laws (where available) is another widely applicable strategy. Citizen wellbeing surveys likewise serve as politically salient outcome indicators that resonate with both decision-makers and the public, particularly in contexts where technical data lacks emotional resonance.

When it comes to the choice of indicators and reporting content, several strategies are purpose-dependent or context-specific. For maximum instrumental use—such as guiding municipal decision-making—aligning reports with local government competencies, budget cycles, and election calendars is especially important. A central design dilemma concerns the tension between sustainability as a long-term, global goal and the need for reports to reflect short-term, local achievements that maintain public and political attention. On one hand, outcome indicators are essential to monitor a city's long-term social and environmental trajectory. On the other hand, output and policy indicators—as promoted by the Dutch LSM and Latin American performance accountability systems (cf. Chapters 4 and 5)—are more easily aligned with local government responsibilities but risk fostering a narrow, short-term orientation. Meanwhile, process indicators can promote the spread of novel sustainability policies but may shift focus from actual results

to effort alone, potentially undercutting substantive progress. This trade-off can be particularly problematic when output indicators dominate and obscure stagnation or decline in sustainability outcomes.

Several contextual factors influence the relevance and viability of these approaches. These include whether municipalities are legally responsible for key policy areas, the length of mayoral terms, and the availability of reliable and disaggregated data. Additionally, data literacy levels and public understanding of sustainability vary widely by region, influencing both indicator selection and report accessibility.

Regarding accountability for government action and goal achievement, performance pressure can indeed help drive results and enhance citizen trust (cf. Chapter 4), particularly in administrative cultures prone to inertia. However, as noted in Table 6.5, such pressure may also produce rigidity or strategic behaviour (gaming), particularly where targets are politically contested or difficult to adapt. These dynamics make performance accountability mechanisms purpose-dependent and potentially incompatible with multi-stakeholder collaboration, especially in politicised settings, where commitments may be diluted or avoided altogether.

In terms of political use and agenda-setting, gaining public attention is a prerequisite. Reporting tends to be most impactful where sustainability information is scarce (filling an information vacuum) or where reports introduce novel evidence (challenging the status quo). As suggested by Chapters 2 and 3, the contribution of sustainability reporting to organisational capacity tends to be more pronounced in contexts with weaker baseline governance capacity, whereas in high-functioning bureaucracies, the added value of reporting may diminish over time.

With respect to report periodicity and timing, effectiveness depends on the report's ability to reach and engage its target audiences. In fast-changing urban contexts—such as many Latin American cities—annual reporting may carry more news value than in more stable European environments. Some cities (cf. Chapter 2) successfully combined regular indicator updates with rotating thematic foci, which helped sustain attention and relevance. Aligning the publication of reports with electoral cycles or major public events, and disseminating through multiple communication channels, are also generally effective tactics.

However, the specific outreach formats remain context-dependent. For instance, Latin American cities (Chapters 3 and 5) have successfully used a combination of press releases, public meetings, and neighbourhood forums to broaden participation. Outreach activities such as roundtables and workshops involving journalists, civil servants, or activists are generally perceived positively; however, in initiatives focused primarily on government accountability, clear boundaries and role separations may be preferable to safeguard impartiality and legitimacy.

6.2.3 Theoretical synthesis: Configuration, temporality, and ecological fit

As emphasised already, the strategies summarised in Table 6.5 do not constitute universal prescriptions or a list of 'best practices'. Moreover, several strategies are interlinked and may not function effectively in isolation. Fundamentally, this dissertation reconceptualizes sustainability reporting not merely as a technical measurement exercise, but as a distinct *mode of urban governance*. Reports do not passively describe urban realities; they actively structure how sustainability is defined and acted upon within specific political contexts. In this view, reporting functions simultaneously as a *conceptual lens* (shaping interpretations), an *organisational mechanism* (coordinating routines), and a *political symbol* (signalling legitimacy). Its effects, therefore, cannot be reduced to merely *supplying data*, but emerge as an *evolving governance practice*.

Synthesizing the findings through this lens, it becomes clear that success depends on three dimensions: internal configuration, temporal dynamics, and ecological fit. First, design choices must work together. A case in point is the Dutch LSM (see Chapter 4), which centred around the combination of voluntary self-assessments and public league tables ranking high- and low-performing municipalities. In the particular political regime of the Netherlands—characterised by high openness and professional autonomy among public servants—this configuration worked well using policy indicators (effort), but would likely not have functioned with outcome indicators (results). The corollary of this finding is that, rather than designing and evaluating sustainability reporting in terms of isolated strategies, it is necessary to consider configurations of interdependent design choices.

Second, these configurations change over time. The empirical studies have shown that although many types of positive effects are achievable in principle, they rarely appear concurrently. On the contrary, there is evidence that some effects compete with one another—much like 'rivalrous' goods in economics. For example, as described in Chapters 2, 4, and 5, the more a reporting system is aligned with a municipality's legal competencies and planning cycles, the more likely it is to yield instrumental use by councillors. Yet this same alignment often makes the report less engaging or comprehensible to citizens, thus diminishing conceptual influence. The divergent informational needs of councillors and citizens as target audiences differ too much for a single format or strategy to serve both equally well. Strategies that work in the initiation phase often lose effectiveness later. For instance, voluntary reporting fosters learning early on but risks becoming a routine exercise without political bite later. Thus, initiatives need planned adaptation to avoid 'reporting fatigue.'

Third, and perhaps most importantly, design choices must be aligned with institutional context. Viewed through the lens of informational governance, reporting initiatives occupy a specific 'ecological niche'. They succeed not when they copy

the existing political culture, but when they provide a resource that is missing. This leads to a 'paradox of context': effective reporting often acts to counter-balance the dominant political features of a jurisdiction.

In the adversarial contexts of Latin America, the main gap is trusted information. Here, the context-specific variable of trust dictates the design. Successful initiatives (Chapter 3) served as neutral platforms to bridge polarised actors. This helps explain the puzzle of why civil society monitoring is less prominent in higher-income nations like Chile and Uruguay: because the state already provides relatively reliable data, there is no empty niche for an external watchdog to fill.

In the consensual democracies of Northern Europe, the gap is rarely data reliability, but rather urgency or coordination. Here, the trade-off shifts. In the Netherlands, the Local Sustainability Meter succeeded by introducing artificial competition to speed up decision-making. In cities like Zurich and Freiburg, reporting served to break down administrative silos. Thus, the relative merit of a strategy depends on its ability to provide the specific type of governance pressure—whether factual evidence or competitive ranking—that the state currently cannot, or will not, provide.

6.3 Contributions, limitations, and research recommendations

As explained in the introduction, city-oriented sustainability reporting can be considered a 'transparency and accountability tool' within the broader field of (urban) sustainability governance. It is therefore appropriate to reflect briefly on this dissertation's contributions to the field, its limitations, and potential directions for further research.

In the last two decades, the body of literature on sustainability governance has increased substantially. Research on sustainability assessment, indicators, and reporting is vast and rapidly growing (Ramos, 2019), covering international frameworks, states, cities, neighbourhoods, value chains, and consumers. In the footsteps of the 'audit explosion' discussed since the 1990s, scholars have also written about an 'indicator explosion' affecting city practices and urban research (Tanguay et al., 2010). However, many publications are either descriptive or prescriptive with limited attention to theory-building and hypothesis testing. In the ever-evolving context of countries and cities, substantial knowledge gaps remain and continue to emerge. In the words of Ramos (2019, p. 824), "despite important progress and the existing vast amount of literature, sustainability indicators are still an underexplored field of study" with a need for a "richer selection of case studies" addressing under-researched areas. The current research project has heeded this call by conducting comparative case studies on underrepresented geographical regions (of Latin America), under-researched topics (such as the use

of policy indicators and accountability mechanisms), and the long-term interplay between the context, design choices, and effects of reporting initiatives.

6.3.1 Contributions to the body of knowledge

One distinctive contribution of this dissertation is its identification of periodicity (e.g., annual vs. biennial reporting) as a key factor shaping the reception and impact of sustainability reports. The research also highlights previously underexamined contextual variables, such as the degree of politicisation. The Latin American case studies revealed that in several countries (e.g., Mexico), civil society groups compiling sustainability reports cooperate effectively with governments, whereas in others—with more contested relations—the local government is purposefully and carefully kept at a distance.

Beyond these empirical observations, this dissertation contributes to the theory of informational governance by identifying the 'ecological niche' of reporting. Challenging the assumption that reporting systems should mirror the administrative culture of their host city (isomorphism), the findings suggest a "paradox of context": effective reporting often acts to counter-balance rather than reflect local political features. As evidenced by the distinction between 'neutral' reporting in adversarial Latin American contexts and 'competitive' benchmarking in the consensual Netherlands, reporting initiatives appear to thrive only when they fill a specific governance deficit that the state cannot, or will not, address.

Another contextual factor identified in this research project is the strength of civil society organisations and "evidence literacy" (Browne, 2022; Pupphachai & Zuidema, 2021) in local communities. This 'niche' hypothesis also helps explain the finding regarding the relative advantages of strict accountability regimes (such as performance reporting for mayors, cf. Chapter 5). These mechanisms depend on the government's baseline effectiveness and trust levels, resonating with much earlier research about New Public Management. In relation to the particular issue of urban sustainability governance, it is in line with the assertion by Mees and Driessen (2019, p. 686) that "striking a balance between accountability and flexibility cannot be easily fixed and will remain a key challenge for interactive governance arrangements and an interesting topic for further empirical studies".

The key finding that sustainability reporting is not a 'jack-of-all-trades' instrument but a niche tool aligns with public management literature that debunks universal claims about governance tools – as Behn (2003) concluded about performance measurement, for example, "different purposes require different measures". Across public management, there is ample evidence of trade-offs and dilemmas, with virtually all innovations potentially bringing benefits yet also drawbacks (Bartke & Schwarze, 2015; Bond & Morrison-Saunders, 2009; Bowen et al., 2017; De Bruijn, 2002). Other scholars generally acknowledge the

importance of temporality in governance research (Sáez, 2012) yet few put it into practice due to the paucity (and high costs) of longitudinal studies.

The observation made in this dissertation that, in numerous countries, urban and regional sustainability are now accepted as a normative discourse, also aligns with extant research. Several researchers refer to bottlenecks in terms of policy implementation and label the ensuing situation as 'paralysis' (Carr, 2013) or 'policy immobility' (McLean & Borén, 2014). In terms of local government action, one important (and underexplored) issue concerns the advantages and drawbacks of integrating sustainability into the local government's strategic planning process (De Matteis & Borgonovi, 2021).

This dissertation has not identified one universal model of 'integrated sustainability planning'. Instead, it offers a more nuanced account: integrated policymaking and reporting may use different indicator types (e.g., outcomes, outputs, policies), invoke distinct accountability mechanisms, and apply differentiated strategies for various stakeholder audiences (e.g., councillors vs. citizens). The latter finding is consistent with the observation of other researchers that "the number of indicators matters depending on whether they are used for external reporting or for internal management reasons" (Manes-Rossi et al., 2020, p. 660).

6.3.2 Limitations

The mapping and case selection (cf. Chapters 2–5) revealed significant variation in the design and stakeholder composition of reporting initiatives across world regions. As summarised in Table 6.6, in Europe most reporting initiatives are organised by local governments and focus on outcome-level indicators. There are exceptions, such as the Dutch and French 'sustainability barometers' (Chapter 4), which focus on local government policies and involve civil society organisations working with government funding. In Latin America, by contrast, many initiatives are run by civil society organisations—typically at arm's length from government involvement. This difference, which can be related to different types of democracies, with European countries being more consensual and Latin America primarily majoritarian (Lijphart, 2012), enriches our understanding of sustainability reporting as an empirical phenomenon. However, this implies a strict limit to generalisability: specific design strategies (e.g., public rankings) are not universally applicable 'best practices' but are contingent on the regime type. A tool that functions as a catalyst in a consensual democracy may act as a destabiliser in a polarised one.

Furthermore, the 'ecological niche' hypothesis proposed in Section 6.2.3 remains an inductive proposition derived from the study of democratic contexts. It is unclear whether the logic of 'counter-balancing' applies to authoritarian or semi-authoritarian regimes (e.g., in parts of Asia or Africa), where the governance deficit is structurally different. Moreover, the selection of 'frontrunners' for the

purpose of longitudinal evaluations implies the risk of a 'survivorship bias' (Section 1.3.4). This means our understanding of the 'niche' is based on successful adaptations; the study lacks data on the 'failed mutations'—initiatives that emerged but could not find a functional role in a given governance ecosystem.

Table 6.6 *City-oriented sustainability reporting systems in Europe and Latin America*

Region	Approach to sustainability reporting	
	Most prevalent stakeholders	Most prevalent focus
Europe	• Governments	• Most initiatives: outcome-level sustainability indicators • Some initiatives: benchmarking of municipal sustainability policies
Latin America	• Governments • Civil society (NGOs, foundations, businesses)	• Most initiatives: outcome-level sustainability indicators • Some initiatives: addressing local government accountability for sustainability policies

As noted in the introductory sections on the conceptual framework (1.3.5) and research methods (1.3.6), another limitation stems from the research's focus on intangible concepts, such as accountability and shifts in public awareness and discourse, which are inherently challenging to define and assess empirically. While the mixed-methods approach and triangulation strategies applied across the four studies were methodologically sound, several findings remain tentative and exploratory. This is both a limitation and an invitation for future research to refine, test, and expand upon the conceptual and empirical insights presented here.

6.3.3 Recommendations for further work

Elmquist et al. (2019) assert that current urban sustainability discourses are filled with misconceptions and vague interpretations that often lead to unintended consequences, hindering progress and concerted action. Others are less scathing in their analysis and argue that in emerging fields of study—such as smart or sustainable cities—one can often observe historical patterns in the research literature, with "prescriptive trends as the first wave, critical papers as the second wave and empirical papers as the third wave—in other words, optimism, then pessimism and, finally, realism" (Grossi et al., 2020, p. 637). To further the understanding of sustainability reporting, additional 'realistic' research is required along three key avenues: i) expanding geographical coverage and testing the 'niche' hypothesis in non-democratic contexts; ii) employing more sophisticated

conceptual and methodological frameworks; and iii) exploring additional aspects of sustainability reporting and urban sustainability governance.

i. Replication studies with enlarged geographical scope

In terms of geographical scope, this research project makes novel contributions by comparing cases from Europe and Latin America. However, the evidence base is limited to democratic settings.It remains an open question whether the 'ecological niche' theory—specifically the need to counter-balance local political culture—holds true in authoritarian or semi-authoritarian regimes. This would be particularly worthwhile through additional case studies—as Mendonça and Grandé (2023, p. 907) argue, "qualitative, case-based studies are important in their own right, as they can detail experiences in the region that could not be conveyed in other ways". Comparative case studies in new regions would also help to assess whether observed effects are context-specific or more widely generalisable. Future research should test whether the logic of 'neutrality vs. friction' applies in contexts where the state holds a monopoly on information, such as in parts of Asia or Africa.

ii. Methodological refinement using an enhanced conceptual framework

The conceptual framework and methods developed in this dissertation lend themselves well to replication. Drawing on findings from the four empirical studies, the original evaluation framework (see Figure 1.7 and Figure 2.2) has been refined to include additional factors, as illustrated in Figure 6.2. For example, the presence of other 'transparency and accountability tools'—such as high-profile sustainability rankings—can shape whether a local reporting initiative gains traction (cf. Chapter 4). Internal governance features, such as membership in networks of like-minded organisations (cf. Chapter 3), also affect continuity and legitimacy. Content-related design choices, including indicator selection, timing relative to electoral cycles, and dissemination methods, are equally critical. The refined framework offers a transferable model that other researchers can use to analyse reporting initiatives in diverse urban contexts.

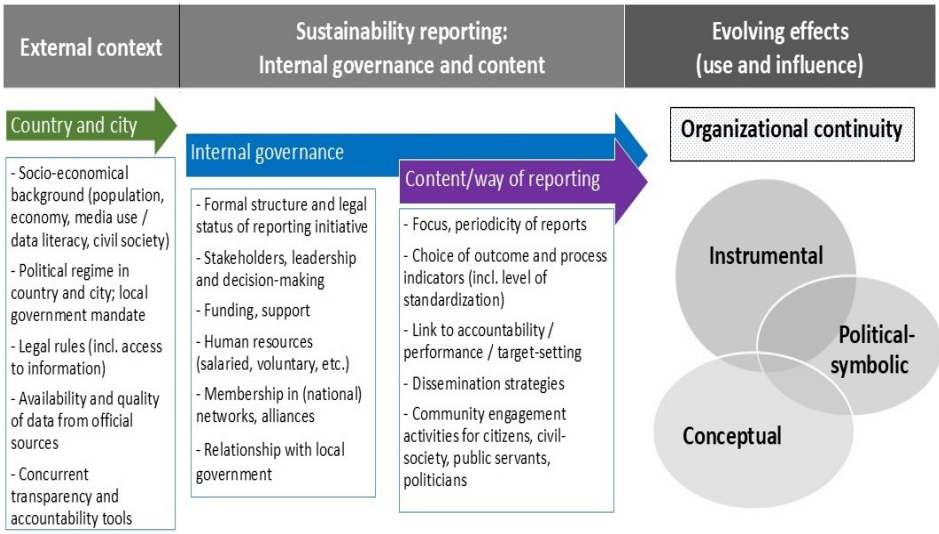


Figure 6.2 *Evaluation framework on the effects of sustainability reporting in divergent contexts*

It is important to note that the conceptual framework used in this research assumes the presence of democratic ideals, which, despite global threats, remain prevalent; recent estimates indicate that "almost half of the world's population live in a democracy of some sort" (Economist Intelligence Unit, 2024).

While studying frontrunners is a well-established method for identifying enabling factors, it introduces limitations—especially when used in longitudinal designs—due to the risk of 'survivorship bias.' To address this, future research could combine longitudinal depth with broader coverage. A promising direction would be a long-term evaluation of sustainability reporting practices across a randomly selected sample of cities, assessed through quantitative analysis or Qualitative Comparative Analysis (QCA) (Schneider & Wagemann, 2012) for variables such as design choices, observed effects, and continuity over time.

iii. Enlarged scope of reporting and sustainability governance

A third major avenue for future research involves expanding the conceptual scope of sustainability reporting. Several themes that emerged across the empirical studies merit deeper investigation. First, the notion of reporting growth—its phases, transitions, and 'maturity'—could be conceptualised more clearly through the lens of 'temporality' identified in Section 6.2.3, drawing on existing frameworks (Lyytimäki, 2019; Rosvall et al., 2023). Future research should investigate how initiatives can transition from data creation (Phase 1) to management integration (Phase 2) without succumbing to 'reporting fatigue'. This would enable

more nuanced assessments of when and how reporting systems evolve or plateau over time. Second, organisational and leadership capacities deserve more systematic study. Findings from this dissertation suggest that municipal leadership, managerial capacities (Ferreira & Botero, 2020; Krantz & Gustafsson, 2021) as well as human resource management (Bowen et al., 2017) significantly influence the design and sustainability reporting initiatives. Similarly, the role of networks—among local governments and civil society—warrants attention as both a macro-level enabler and constraint (Fenton & Gustafsson, 2017). Third, future research should better integrate citizen behaviour and perception, bridging macro frameworks with micro-level drivers. As Wiedman and Allen (2021, p. 1) argue, "SDG monitoring and assessment of cities should take advantage of both consumption-based (footprint) accounting and benchmarking against planetary boundaries and social thresholds in relative and absolute terms". This is particularly relevant for cities, whose sustainability depends on global flows of resources and services. Theoretical models from psychology and behavioural sciences (Lewis, 2015) could enrich this line of inquiry. Finally, technology is reshaping sustainability governance. Smart city platforms (Grossi et al., 2020) and digital democracy tools (Helbing et al., 2023) increasingly influence both data availability and civic engagement, making them vital topics for future research.

6.4 Implications for policy and practice

For decades, policymaking in the fields of transparency and accountability has been dominated by simplistic ideas about information use. There is the "rationalist assumption that if you give people data, they will make good use of it" (Moreno Pires et al., 2017, p. 11). This belief also underlies some activist-led reporting initiatives that expect immediate policy impacts. As a result, sustainability reporting is sometimes promoted as a universal solution—capable of simultaneously improving decision-making and citizen engagement.

In a sense, such a 'sales pitch' is legitimate as sustainability reporting has the potential of benefits in multiple domains. Furthermore, whenever a new 'accountability and transparency tool' (or any other tool, for that matter) is introduced, some excited anticipation of positive results may help bring these about as a self-fulfilling prophecy. If the general expectation is negative, people and institutions will probably not invest time and energy, thus robbing the tool of its potential transformative power.

However, unrealistic expectations are likely to backfire. As this dissertation showed, numerous sustainability reporting initiatives were abandoned when they did not produce the transformative impacts that some of their creators had–perhaps naively–hoped for. Therefore, a realistic, evidence-informed understanding of the conditions, costs, and benefits of reporting (and its specific 'ecological niche') is essential for long-term viability.

The costs of sustainability reporting vary widely. They depend primarily on (i) the availability of reliable data and (ii) key design decisions—such as the number and type of indicators, frequency of reporting, and extent of consultation. One important but underappreciated benefit is institutional learning. Even when initiatives lack formally stated objectives, they often spark positive internal and inter-organisational exchanges. Notably, none of the initiatives analysed—despite preambles or official aims—explicitly included learning as a goal, yet it consistently emerged as a beneficial side effect. Another critical insight is that reporting initiatives do not operate in a vacuum: their success can be constrained by the presence of competing tools, particularly when these involve benchmarking or require staff engagement. As demonstrated in several cases, overlapping instruments may divide attention and diminish the visibility or legitimacy of any single initiative.

Based on the synthesis of empirical and conceptual findings presented in this dissertation, five interlinked domains are particularly relevant for practitioners involved in the initiation, design, or support of sustainability reporting systems. These domains provide a structured set of considerations to inform evidence-based policy and practice and are summarised in Box 6.1.

Box 6.1. *Five Key Domains for Designing Effective Sustainability Reporting Initiatives in Cities*

1. Understand the Local Context

Every city operates within a unique constellation of institutional capacities, political dynamics, civil society engagement, and data infrastructure. Before launching a reporting initiative, these conditions should be systematically mapped. Key considerations include:

- The level of politicisation and trust in government: in high-trust, consensual political systems, collaborative reporting efforts between governments and civil society actors may strengthen legitimacy and outcomes Here, reporting can serve to introduce necessary friction (e.g., via benchmarking). . In more adversarial or majoritarian settings (e.g., parts of Latin America), clear institutional boundaries may be necessary for independence and credibility to provide a "safe harbour" of neutral data.
- The availability of reliable data, and the city's institutional ability to produce and interpret indicators.
- The presence of competing transparency tools, which may limit the public visibility or administrative uptake of any single initiative.

2. Clarify Purpose and Anticipated Effects

Sustainability reporting can serve multiple functions—instrumental, conceptual, and political-symbolic—but no single initiative can fulfil all of these equally well. Trade-offs between these effects are common. Practitioners should be explicit about which form(s) of use are prioritised, recognising that benefits may evolve over time. Clarity of purpose is essential:

- For instrumental use, align indicators and reporting cycles with existing policy mandates, budget processes, and legal responsibilities.
- For conceptual use, emphasise organisational learning through stakeholder dialogue and inter-institutional exchange (typically most effective in the early phases of an initiative).
- For political-symbolic effects, select indicators with high salience for the public and design communication strategies to foster visibility and legitimacy.

3. Choose Governance and Organisational Design

The internal governance of a reporting initiative significantly influences its continuity and credibility. Arrangements should be context-sensitive and adapted to the local political culture and institutional landscape. Key design choices include:

- The degree of centralisation: reporting can be coordinated by a government unit, a civil society network, or a hybrid structure. Each model involves trade-offs between control, legitimacy, and stakeholder buy-in.
- The participation model: voluntary participation can foster learning and innovation but may lead to discontinuity; legal mandates enhance consistency but risk strategic behaviour or 'gaming' of indicators.
- The scope of stakeholder involvement: broader alliances (e.g., involving media or academia) can enhance credibility and outreach and provide protection in polarised contexts but may complicate coordination and dilute accountability.

4. Structure Reporting Content and Communication

Design choices around indicators, timing, formats, and dissemination strategies shape the relevance and impact of the reporting system. Recommendations include:

- Use standardised indicator sets for comparability but complement them with context-specific indicators relevant to local sustainability challenges.

- Combine outcome indicators with policy and process indicators to address both long-term trends and short-term action.
- Avoid information overload: a manageable number of meaningful indicators is preferable to large, overly technical sets.
- Periodicity matters: empirical cases show that annual reporting can foster learning and visibility, especially in fast-changing environments, while biennial or triennial reports may be more feasible in resource-constrained contexts.
- Consider alternating focus themes across reporting cycles (e.g., air quality in one year, housing equity in the next), which can sustain interest and and mitigate reporting fatigue..
- Time report releases to coincide with electoral cycles, policy events, or budget processes, and disseminate findings through multi-channel strategies, including dashboards, press events, and local fora.

5. Embed Evaluation and Adaptability

Sustainability reporting is not a one-off exercise but a dynamic process that must evolve with changing conditions. Reporting systems that are flexible, reflexive, and adaptive are more likely to retain legitimacy and effectiveness over the long term. To ensure continued relevance and credibility:

- Implement regular self-evaluation, including stakeholder feedback and assessment of effects.
- Monitor shifts in the external environment, including emerging tools or changes in public discourse.
- Be prepared to adapt design features—such as indicator sets, partnerships, or communication strategies—based on lessons learned over time to ensure the initiative continues to fill a relevant niche in the governance ecosystem.

The key practical takeaway of this dissertation is that sustainability reporting should not be treated as a universal solution. Rather than aiming for an all-purpose instrument, designers should develop initiatives that reflect the specific trade-offs, capacities, and objectives relevant to their context. Tools designed without such clarity risk becoming 'jack-of-all-trades but master-of-none'—failing to gain traction or deliver impact. In contrast, initiatives that are consciously designed, context-sensitive, and periodically adapted can make significant contributions to urban sustainability at relatively low cost.

7. References

Acuto, M., Pejic, D., & Briggs, J. (2021). Taking City Rankings Seriously: Engaging with Benchmarking Practices in Global Urbanism. *International Journal of Urban and Regional Research*, 1468-2427.12974. https://doi.org/10.1111/1468-2427.12974

Adams, C. A. (2015). The International Integrated Reporting Council: A call to action. *Critical Perspectives on Accounting*, *27*, 23-28. https://doi.org/10.1016/j.cpa.2014.07.001

Adams, C. A., & Frost, G. R. (2008). Integrating sustainability reporting into management practices. *Accounting Forum*, *32*(4), 288-302. https://doi.org/10.1016/j.accfor.2008.05.002

Adams, S., & Simnett, R. (2011). Integrated Reporting: An Opportunity for Australia's Not-for-Profit Sector: Integrated Reporting for Not-for-Profit Sector. *Australian Accounting Review*, *21*(3), 292-301. https://doi.org/10.1111/j.1835-2561.2011.00143.x

Adaptation. (2024). In *Cambridge University Press*. https://dictionary.cambridge.org/dictionary/english/adaptation

Alcaraz-Quiles, F. J., Navarro-Galera, A., & Ortiz-Rodríguez, D. (2014). Factors determining online sustainability reporting by local governments. *International Review of Administrative Sciences*, 0020852314541564. https://doi.org/10.1177/0020852314541564

Ammons, D. N., & Madej, P. M. (2017). Citizen-Assisted Performance Measurement? Reassessing Its Viability and Impact. *The American Review of Public Administration*, 0275074017713295. https://doi.org/10.1177/0275074017713295

Andrews, M. (2010). Good government means different things in different countries. *Governance*, *23*(1), 7-35.

Andrews, M., Pritchett, L., & Woolcock, M. (2017). *Building state capability: Evidence, analysis, action*. http://dx.doi.org/10.1093/acprof:oso/9780198747482.001.0001

Aust, H. P., & Plessis, A. du. (2018). Good urban governance as a global aspiration: On the potential and limits of SDG 11. In D. French & L. Kotzé (Red.), *Sustainable Development Goals. Law, Theory and Implementation* (pp. 201-221). Edward Elgar Publishing. https://www.elgaronline.com/edcollchap/edcoll/9781786438751/9781786438751.00016.xml

Baliña, L. M. (2017). *Impactos institucionales, reformas organizacionales y nuevas prácticas a partir de la implementación de Programas y Planes de Metas en las ciudades de San Pablo y Córdoba entre 2008 y 2016* [Tesis de Maestría, Universidad Católica de Córdoba]. http://pa.bibdigital.ucc.edu.ar/1543/

Ballantine, J. (2014, december 1). The value of sustainability reporting by cities. *Cities Today*. http://cities-today.com/2014/12/value-sustainability-reporting-cities/

Barrington-Leigh, C. (2017). The Role of Subjective Well-Being as an Organizing Concept for Community Indicators. In M. Holden, R. Phillips, & C. Stevens (Red.), *Community Quality-of-Life Indicators: Best Cases VII* (pp. 19-33). Springer.

Barrington-Leigh, C., & Escande, A. (2016). Measuring Progress and Well-Being: A Comparative Review of Indicators. *Social Indicators Research*. https://doi.org/10.1007/s11205-016-1505-0

Bartke, S., & Schwarze, R. (2015). No perfect tools: Trade-offs of sustainability principles and user requirements in designing support tools for land-use decisions between greenfields and brownfields. *Journal of Environmental Management*, *153*, 11-24. https://doi.org/10.1016/j.jenvman.2015.01.040

Bartocci, L., & Picciaia, F. (2013). Towards Integrated Reporting in the Public Sector. In C. Busco, M. Frigo, A. Riccaboni, & P. Quattrone (Red.), *Integrated Reporting. Concepts and Cases that Redefine Corporate Accountability* (pp. 191-204). Springer. http://link.springer.com/chapter/10.1007/978-3-319-02168-3_12

Batley, R., & Mcloughlin, C. (2015). The Politics of Public Services: A Service Characteristics Approach. *World Development*, *74*, 275-285. https://doi.org/10.1016/j.worlddev.2015.05.018

Bauler, T. (2012). An analytical framework to discuss the usability of (environmental) indicators for policy. *Ecological indicators*, *17*, 38-45.

Baumüller, J., Schaffhauser-Linzatti, M. M., & Sopp, K. (2024). An Agenda for Research on Sustainability Reporting: Europe and Beyond. In W. Leal Filho, A. L. Salvia, & C. R. Portela De Vasconcelos (Red.), *An Agenda for Sustainable Development Research* (pp. 635-650). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-65909-6

Bebbington, J., Higgins, C., & Frame, B. (2009). Initiating sustainable development reporting: Evidence from New Zealand. *Accounting, Auditing & Accountability Journal*, *22*(4), 588-625. https://doi.org/10.1108/09513570910955452

Bebbington, J., & Larrinaga, C. (2014). Accounting and sustainable development: An exploration. *Accounting, Organizations and Society*, *39*(6), 395-413.

Bebbington, J., & Unerman, J. (2018). Achieving the United Nations Sustainable Development Goals: An enabling role for accounting research. *Accounting, Auditing & Accountability Journal*, *31*(1), 2-24. https://doi.org/10.1108/AAAJ-05-2017-2929

Becker, T. (2015). Mobilising sustainability: Why European sustainable urban development initiatives are slow to materialise. *Planning Theory & Practice*, *16*(1), 103-108. https://doi.org/10.1080/14649357.2014.991544

Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public administration review*, *63*(5), 586-606.

Bell, S., & Morse, S. (Red.). (2018a). *Routledge Handbook of Sustainability Indicators*. Routledge. https://doi.org/10.4324/9781315561103

Bell, S., & Morse, S. (2018b). Sustainability Indicators Past and Present: What Next? *Sustainability*, *10*(5), Article 5. https://doi.org/10.3390/su10051688

Bexell, M., & Jönsson, K. (2019). Country Reporting on the Sustainable Development Goals—The Politics of Performance Review at the Global-National Nexus. *Journal of Human Development and Capabilities*, *20*(4), 403-417. https://doi.org/10.1080/19452829.2018.1544544

Biermann, F., Hickmann, T., Sénit, C.-A., Beisheim, M., Bernstein, S., Chasek, P., Grob, L., Kim, R. E., Kotzé, L. J., & Nilsson, M. (2022). Scientific evidence on the political impact of the Sustainable Development Goals. *Nature Sustainability*, *5*(9), 795-800.

Biermann, F., Kanie, N., & Kim, R. E. (2017). Global governance by goal-setting: The novel approach of the UN Sustainable Development Goals. *Current Opinion in Environmental Sustainability*, *26-27*, 26-31. https://doi.org/10.1016/j.cosust.2017.01.010

Biondi, L., & Bracci, E. (2018). Sustainability, Popular and Integrated Reporting in the Public Sector: A Fad and Fashion Perspective. *Sustainability*, *10*(9), Article 9. https://doi.org/10.3390/su10093112

Boaz, A., & Nutley, S. (2015). Evidence-based policy and practice. In T. Bovaird & E. Löffler (Red.), *Public Management and Governance* (Third edition, pp. 376-392). Taylor & Francis.

Bond, A. J., & Morrison-Saunders, A. (2009). Sustainability appraisal: Jack of all trades, master of none? *Impact assessment and project appraisal*, *27*(4), 321-329.

Bose, S. (2020). Evolution of ESG Reporting Frameworks. In D. C. Esty & T. Cort (Red.), *Values at Work* (pp. 13-33). Springer International Publishing. https://doi.org/10.1007/978-3-030-55613-6_2

Bouckaert, G., & van Dooren, W. (2015). Performance measurement and management in public sector organizations. In T. Bovaird & E. Löffler (Red.), *Public Management and Governance* (Third edition, pp. 148-161). Taylor & Francis.

Bovens, M. (2005). Public Accountability. In E. Ferlie, L. Lynn, & C. Pollitt (Red.), *The Oxford Handbook of Public Management* (pp. 182-208). Oxford University Press.

Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal*, *13*(4), 447-468.

Bovens, M., Schillemans, T., & Hart, P. 'T. (2008). Does Public Accountability Work? An Assessment Tool. *Public Administration*, *86*(1), 225-242. https://doi.org/10.1111/j.1467-9299.2008.00716.x

Bowen, K. J., Cradock-Henry, N. A., Koch, F., Patterson, J., Häyhä, T., Vogt, J., & Barbi, F. (2017). Implementing the "Sustainable Development Goals": Towards addressing three key governance challenges—collective action, trade-offs, and accountability. *Current Opinion in Environmental Sustainability*, *26-27*, 90-96. https://doi.org/10.1016/j.cosust.2017.05.002

Brandsma, G. J., & Schillemans, T. (2012). The Accountability Cube: Measuring Accountability. *Journal of Public Administration Research and Theory*, *23*(4), 953-975. https://doi.org/10.1093/jopart/mus034

Bravo, N. (2019). Voto Programático: La condición previa a las elecciones territoriales de 2020. In H. von Baer & N. Bravo (Red.), *Desarrollo territorial colaborativo: Descentralizando poder, competencias y recursos* (pp. 741-764). Ediciones Universidad de La Frontera.

Brown, H. S., de Jong, M., & Levy, D. L. (2009). Building institutions based on information disclosure: Lessons from GRI's sustainability reporting. *Journal of Cleaner Production*, *17*(6), 571-580. https://doi.org/10.1016/j.jclepro.2008.12.009

Browne, G. R. (2022). An evidence cycle framework for community development initiatives. *Community Development Journal*, 19. https://doi.org/10.1093/cdj/bsac011

Cabrera Paredes, M., González Moscoso, E., & Vintimilla Carrasco, M. I. (2020). *Colectivo Cuenca Ciudad Para Vivir. Diez años de apuesta para la ciudanía activa. Sistematización de la experiencia*. CCCV.

Cáceres, P. (2014). Planes y Programas de Metas como innovaciones en los procesos de rendición de cuentas en el nivel local. *Pensamiento Proprio*, *19*(40), 191-226.

Cairney, P., & Oliver, K. (2017). Evidence-based policymaking is not like evidence-based medicine, so how far should you go to bridge the divide between evidence and policy? *Health Research Policy and Systems*, *15*(1). https://doi.org/10.1186/s12961-017-0192-x

Caprotti, F., Cowley, R., Datta, A., Broto, V. C., Gao, E., Georgeson, L., Herrick, C., Odendaal, N., & Joss, S. (2017). The New Urban Agenda: Key opportunities and challenges for policy and practice. *Urban Research & Practice*, *10*(3), 367-378. https://doi.org/10.1080/17535069.2016.1275618

Cardoni, A., & Kiseleva, E. (2023). *Sustainable Governance: Concept, Metrics and Contexts*. Springer International Publishing. https://doi.org/10.1007/978-3-031-37492-0

Carnohan, S. A., Trier, X., Liu, S., Clausen, L. P. W., Clifford-Holmes, J. K., Hansen, S. F., Benini, L., & McKnight, U. S. (2023). Next generation application of DPSIR for sustainable policy implementation. *Current Research in Environmental Sustainability*, *5*, 100201. https://doi.org/10.1016/j.crsust.2022.100201

Carr, C. (2013). Discourse Yes, Implementation Maybe: An Immobility and Paralysis of Sustainable Development Policy. *European Planning Studies*, *22*(9), 1824-1840. https://doi.org/10.1080/09654313.2013.806433

Castellaci, F., Grodal, S., Mendonca, S., & Wibe, M. (2005). Advances and challenges in innovation studies. *Journal of Economic Issues*, *39*(1), 91-121.

Ceulemans, K., Lozano, R., & Alonso-Almeida, M. (2015). Sustainability Reporting in Higher Education: Interconnecting the Reporting Process and Organisational Change Management for Sustainability. *Sustainability*, *7*(7), 8881-8903. https://doi.org/10.3390/su7078881

Cidade de Sao Paulo. (2022). *Cartilha de Replicabilidade. Programa de Metas de Sao Paulo*.

City of Amsterdam. (2015). *Sustainable Amsterdam. Agenda for renewable energy, clear air, a circular economy and a climate-resilient city* [Adopted by the Municipal Council of Amsterdam]. City of Amsterdam.

Coenen, F. (2008). The Local Level: New Interpretations of Governance for Sustainable Development in the. Netherlands. In S. Baker & K. Eckerberg (Red.), *New Governance Practices at the Sub-national Level in Europe* (pp. 190-2007). Routledge.

Cohen, M. (2017). A Systematic Review of Urban Sustainability Assessment Literature. *Sustainability*, *9*(11), 2048. https://doi.org/10.3390/su9112048

Cohen, M., Wiek, A., Kay, B., & Harlow, J. (2015). Aligning Public Participation to Stakeholders' Sustainability Literacy—A Case Study on Sustainable Urban Development in Phoenix, Arizona. *Sustainability*, *7*(7), 8709-8728. https://doi.org/10.3390/su7078709

Cohen, S., & Karatzimas, S. (2014). Tracing the Future of Reporting in the Public Sector: Introducing Integrated Popular Reporting. *Available at SSRN 2524807*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2524807

Commissariat Général au Développement Durable. (2012). *Premiers éléments méthodologiques pour l'élaboration du rapport sur la situation en matière de développement durable à l'usage des collectivités territoriales et EPCI à fiscalité propre de plus de 50 000 habitants*. Service de l'Économie, de l'Évaluation et de l'Intégration du Développement Durable.

Correia, C. R., & Roseland, M. (2022). Addressing Negative Externalities of Urban Development: Toward a More Sustainable Approach. *Urban Science*, *6*(2), Article 2. https://doi.org/10.3390/urbansci6020038

Cox, D., Frere, M., & West, S. (2010). Developing and Using Local Community Wellbeing Indicators: Learning from the Experience of Community Indicators Victoria. *Australian Journal of Social Issues*, *45*(1), 71-88.

Cruz Pérez, I. E. (2024). Regulación legal sobre programas electorales en 18 países de América Latina. *Apuntes Electorales*, *23*, 107-138. https://doi.org/10.53985/ae.v23i70.907

Dahl, A. L. (2012). Achievements and gaps in indicators for sustainability. *Ecological Indicators*, *17*, 14-19.

Dahler-Larsen, P. (2013). Constitutive Effects of Performance Indicators: Getting beyond unintended consequences. *Public Management Review*, *16*(7), 969-986. https://doi.org/10.1080/14719037.2013.770058

Davern, M. T., Gunn, L., Giles-Corti, B., & David, S. (2016). Best Practice Principles for Community Indicator Systems and a Case Study Analysis: How Community Indicators Victoria is Creating Impact and Bridging Policy, Practice and Research. *Social Indicators Research*, *131*(2), 567-586. https://doi.org/10.1007/s11205-016-1259-8

De Bruijn, H. (2002). Performance measurement in the public sector: Strategies to cope with the risks of performance measurement. *International Journal of Public Sector Management*, *15*(7), 578-594.

de Jong, M., Joss, S., Schraven, D., Zhan, C., & Weijnen, M. (2015). Sustainable–Smart–Resilient–Low Carbon–Eco–Knowledge Cities; Making sense of a multitude of concepts promoting sustainable urbanization. *Journal of Cleaner Production*, *109*, 25-38. https://doi.org/10.1016/j.jclepro.2015.02.004

de Lancer Julnes, P. (2006). Performance measurement: An effective tool for government accountability? The debate goes on. *Evaluation*, *12*(2), 219.

de Lancer Julnes, P., Broom, C., & Park, S. (2019). A Suggested Model for Integrating Community Indicators with Performance Measurement. Challenges and Opportunities. *International Journal of Community Well-Being*. https://doi.org/10.1007/s42413-019-00046-6

De Matteis, F., & Borgonovi, E. (2021). A Sustainability Management Model for Local Government: An Explanatory Study. *Administrative Sciences*, *11*(4), 126. https://doi.org/10.3390/admsci11040126

de Villiers, C., Dimes, R., & Molinari, M. (2024). How will AI text generation and processing impact sustainability reporting? Critical analysis, a conceptual framework and avenues for future research. *Sustainability Accounting, Management and Policy Journal*, *15*(1), 96-118.

de Villiers, C., Dumay, J., Farneti, F., Jia, J., & Li, Z. (2025). Reprint of: Does mandating corporate social and environmental disclosure improve social and environmental performance?: Broad-based evidence regarding the effectiveness of Directive 2014/95/EU. *The British Accounting Review*, *57*(1), 101558.

de Villiers, C., Rinaldi, L., & Unerman, J. (2014). Integrated Reporting: Insights, gaps and an agenda for future research. *Accounting, Auditing & Accountability Journal*, *27*(7), 1042-1067. https://doi.org/10.1108/AAAJ-06-2014-1736

De Vries, M. S. (2010). Performance measurement and the search for best practices. *International Review of Administrative Sciences*, *76*(2), 313-330.

Del Río Castro, G., González Fernández, M. C., & Uruburu Colsa, Á. (2021). Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): A holistic review. *Journal of Cleaner Production*, *280*, 122204. https://doi.org/10.1016/j.jclepro.2020.122204

Dhakal, S., & Imura, H. (2003). Policy-based indicator systems: Emerging debates and lessons. *Local Environment*, *8*(1). http://www.tandfonline.com/doi/pdf/10.1080/13549830306674

Díaz Jiménez, O., & Natal, A. (2014). *Observatorios ciudadanos: Nuevas formas de participación de la sociedad*. Ediciones Gernika. https://doi.org/10.13140/RG.2.1.2390.9284

Downe, J., Grace, C., Martin, S., & Nutley, S. (2010). Theories Of Public Service Improvement. *Public Management Review*, *12*(5), 663-678. https://doi.org/10.1080/14719031003633201

Dumay, J., Guthrie, J., & Farneti, F. (2010). GRI Sustainability Reporting Guidelines For Public And Third Sector Organizations. A critical review. *Public Management Review*, *12*(4), 531-548. https://doi.org/10.1080/14719037.2010.496266

Ebrahim, A. (2005). Accountability Myopia: Losing Sight of Organizational Learning. *Nonprofit and Voluntary Sector Quarterly*, *34*(1), 56-87. https://doi.org/10.1177/0899764004269430

Economist Intelligence Unit. (2024). *Democracy Index 2023*. https://www.eiu.com/n/campaigns/democracy-index-2023/

Edelenbos, J., Molenveld, A., Mojanchevska, K., Ensenado, E., Ballinas, M. B.-P., Esteban, A., Ruijsink, S., Igalla, M., & Tsatsou, A. (2021). Community-based initiatives in the urban realm what conditions their performance? *Journal of Environmental Planning and Management*, *64*(9), 1689-1712. https://doi.org/10.1080/09640568.2020.1837088

Ekardt, F. (2024). *Sustainability: Transformation, Governance, Ethics, Law* (Vol. 5). Springer International Publishing. https://doi.org/10.1007/978-3-031-62711-8

Elgert, L. (2016). The double edge of cutting edge: Explaining adoption and non-adoption of the STAR rating system and insights for sustainability indicators. *Ecological Indicators*, *67*, 556-564. https://doi.org/10.1016/j.ecolind.2016.02.051

Elgert, L. (2018). Rating the sustainable city: 'Measurementality', transparency, and unexpected outcomes at the knowledge-policy interface. *Environmental Science & Policy*, *79*, 16-24. https://doi.org/10.1016/j.envsci.2017.10.006

Elmqvist, T., Andersson, E., Frantzeskaki, N., McPhearson, T., Olsson, P., Gaffney, O., Takeuchi, K., & Folke, C. (2019). Sustainability and resilience for transformation in the urban century. *Nature Sustainability*, *2*(4), 267-273. https://doi.org/10.1038/s41893-019-0250-1

Estevez, E., Cenci, K., Fillottrani, P., & Janowski, T. (2021). Review of International Standards and Policy Guidelines for Smart Sustainable Cities. In E. Estevez, T. A. Pardo, & H. J. Scholl (Red.), *Smart Cities and Smart Governance* (Vol. 37, pp. 69-99). Springer International Publishing. https://doi.org/10.1007/978-3-030-61033-3_4

Feinstein, O. N., & Moreno, M. G. (2015). Monitoring and evaluation. In J. Kaufmann, M. Sanginés, & M. García Moreno (Red.), *Building Effective Governments: Achievements and Challenges for Results-Based Public Administration in Latin America and the Caribbean* (pp. 195-224). Inter-American Development Bank. https://doi.org/10.18235/0000016

Fenton, P., & Gustafsson, S. (2017). Moving from high-level words to local action—Governance for urban sustainability in municipalities. *Current Opinion in Environmental Sustainability*, *26-27*, 129-133. https://doi.org/10.1016/j.cosust.2017.07.009

Ferreira da Cruz, N., Rode, P., & McQuarrie, M. (2019). New urban governance: A review of current themes and future priorities. *Journal of Urban Affairs*, *41*(1), 1-19. https://doi.org/10.1080/07352166.2018.1499416

Ferreira, M., & Botero, A. (2020). Experimental governance? The emergence of public sector innovation labs in Latin America. *Policy Design and Practice*, *3*(2), 150-162. https://doi.org/10.1080/25741292.2020.1759761

Fiabane, D. F., Alves, M. A., & de Brelàz, G. (2014). Social Accountability as an Innovative Frame in Civic Action: The Case of Rede Nossa São Paulo. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *25*(3), 849-849. https://doi.org/10.1007/s11266-013-9426-x

Flinders, M. (2011). Daring to be a Daniel: The Pathology of Politicized Accountability in a Monitory Democracy. *Administration & Society*, *43*(5), 595-619. https://doi.org/10.1177/0095399711403899

Flower, J. (2015). The International Integrated Reporting Council: A story of failure. *Critical Perspectives on Accounting*, *27*, 1-17. https://doi.org/10.1016/j.cpa.2014.07.002

Fox, J. A. (2015). Social Accountability: What Does the Evidence Really Say? *World Development*, *72*, 346-361. https://doi.org/10.1016/j.worlddev.2015.03.011

Fraser, E. D., Dougill, A. J., Mabee, W. E., Reed, M., & McAlpine, P. (2006). Bottom up and top down: Analysis of participatory processes for sustainability indicator identification as a pathway to community empowerment and sustainable environmental management. *Journal of environmental management*, *78*(2), 114-127.

Friendly, A. (2017). Urban Policy, Social Movements, and the Right to the City in Brazil. *Latin American Perspectives*, *44*(2), 132-148. https://doi.org/10.1177/0094582X16675572

Fuhr, H., Hickmann, T., & Kern, K. (2018). The role of cities in multi-level climate governance: Local climate policies and the 1.5 °C target. *Current Opinion in Environmental Sustainability*, *30*, 1-6. https://doi.org/10.1016/j.cosust.2017.10.006

Garzillo, C., Rok, A., & Fernández de la Fuente, I. (2014). Building on experience. Learning from end users of tools. In M. Joas, K. Theobald, D. McGuinness, C. Garzillo, & S. Kuhn (Red.), *Informed Cities: Making Research Work for Local Sustainability* (pp. 111-129). Routledge.

Gaspardo, M., & Ferreira, M. (2017). Inovação institucional e democracia participativa: Mapeamento legislativo da Emenda do Programa de Metas. *Revista de Administração Pública*, *51*, 129-146. https://doi.org/10.1590/0034-7612148181

Gerring, J. (2007). *Case study research: Principles and practices*. Cambridge University Press.

Giles-Corti, B., Vernez-Moudon, A., Reis, R., Turrell, G., Dannenberg, A. L., Badland, H., Foster, S., Lowe, M., Sallis, J. F., Stevenson, M., & Owen, N. (2016). City planning and population health: A global challenge. *The Lancet*, *388*(10062), 2912-2924. https://doi.org/10.1016/S0140-6736(16)30066-6

Glass, L.-M., & Newig, J. (2019). Governance for achieving the Sustainable Development Goals: How important are participation, policy coherence,

reflexivity, adaptation and democratic institutions? *Earth System Governance*, *2*, 100031. https://doi.org/10.1016/j.esg.2019.100031

Global Reporting Initiative. (2005). *Sector Supplement for Public Agencies*.

Goswami, K., & Lodhia, S. (2014). Sustainability disclosure patterns of South Australian local councils: A case study. *Public Money & Management*, *34*(4), 273-280. https://doi.org/10.1080/09540962.2014.920200

Grandvoinnet, H., Aslam, G., & Raha, S. (2015). *Opening the Black Box: The Contextual Drivers of Social Accountability*. The World Bank. http://elibrary.worldbank.org/doi/book/10.1596/978-1-4648-0481-6

Gray, R. (2010). Is accounting for sustainability actually accounting for sustainability...and how would we know? An exploration of narratives of organisations and the planet. *Accounting, Organizations and Society*, *35*(1), 47-62. https://doi.org/10.1016/j.aos.2009.04.006

Greco, G., Sciulli, N., & D'onza, G. (2012). From Tuscany to Victoria: Some Determinants of Sustainability Reporting by Local Councils. *Local Government Studies*, *38*(5), 681-705. https://doi.org/10.1080/03003930.2012.679932

Greenwood, T. (2008). Bridging the divide between community indicators and government performance measurement. *National Civic Review*, *97*(1), 55-59. https://doi.org/10.1002/ncr.207

Greiling, D., & Grüb, B. (2014). Sustainability reporting in Austrian and German local public enterprises. *Journal of Economic Policy Reform*, *17*(3), 209-223. https://doi.org/10.1080/17487870.2014.909315

Grimmelikhuijsen, S., & Kasymova, J. (2015). Not So Universal After All: Exploring the Meaning and Use of Government Transparency in Consensual and Majoritarian Democracies. *Public Integrity*, *17*(4), 389-407.

Grönholm, S. (2014). Explorative application of monitoring tools. In M. Joas, K. Theobald, D. McGuinness, C. Garzillo, & S. Kuhn (Red.), *Informed Cities: Making Research Work for Local Sustainability* (pp. 52-61). Routledge.

Grönholm, S., & Berrini, M. (2014). Approaches to monitoring sustainable urban development in Europe. In M. Joas, K. Theobald, D. McGuinness, C. Garzillo, & S. Kuhn (Red.), *Informed Cities: Making Research Work for Local Sustainability* (pp. 20-39). Routledge.

Grossi, G., Meijer, A., & Sargiacomo, M. (2020). A public management perspective on smart cities: 'Urban auditing' for management, governance and accountability. *Public Management Review*, *22*(5), 633-647. https://doi.org/10.1080/14719037.2020.1733056

Guerrero-Gómez, T., Navarro-Galera, A., & Ortiz-Rodríguez, D. (2021). Promoting Online Transparency to Help Achieve the Sustainable Development Goals: An Empirical Study of Local Governments in Latin America. *Sustainability*, *13*(4), 1837. https://doi.org/10.3390/su13041837

Gugerty, M. K. (2009). Signaling virtue: Voluntary accountability programs among nonprofit organizations. *Policy Sciences*, *42*(3), 243-273. https://doi.org/10.1007/s11077-009-9085-3

Guinn, D. E., & Straussman, J. D. (2018). Is best practice in development still viable? The case of financial management in fragile Afghanistan. *Public*

Management Review, *0*(0), 1-18. https://doi.org/10.1080/14719037.2018.1441429

Guthrie, J., Ball, A., & Farneti, F. (2010). Advancing Sustainable Management of Public and Not For Profit Organizations. *Public Management Review*, *12*(4), 449-459. https://doi.org/10.1080/14719037.2010.496254

Hahn, R., & Kühnen, M. (2013). Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production*, *59*, 5-21. https://doi.org/10.1016/j.jclepro.2013.07.005

Hák, T., Janoušková, S., & Moldan, B. (2016). Sustainable Development Goals: A need for relevant indicators. *Ecological Indicators*, *60*, 565-573. https://doi.org/10.1016/j.ecolind.2015.08.003

Halachmi, A. (2014). Accountability Overloads. In M. Bovens, R. Goodin, & T. Schillemans (Red.), *The Oxford Handbook of Public Accountability* (pp. 560-573). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641253.013.0011

Halla, P. J. (2021). *From metrics towards meaningful assessments of urban sustainability—Exploring visions, frameworks and indicators*. École Polytechnique Fédérale de Lausanne.

Halla, P. J., & Merino-Saum, A. (2022). Conceptual frameworks in indicator-based assessments of urban sustainability—An analysis based on 67 initiatives. *Sustainable Development*, sd.2300. https://doi.org/10.1002/sd.2300

Hartley, J., & Downe, J. (2007). The shining lights? Public service awards as an approach to service improvement. *Public Administration*, *85*(2), 329-353.

Hayden, A., & Wilson, J. (2016). Is It What You Measure That Really Matters? The Struggle to Move beyond GDP in Canada. *Sustainability*, *8*(7), 623. https://doi.org/10.3390/su8070623

Helbing, D., Mahajan, S., Fricker, R. H., Musso, A., Hausladen, C. I., Carissimo, C., Carpentras, D., Stockinger, E., Argota Sanchez-Vaquerizo, J., Yang, J. C., Ballandies, M. C., Korecki, M., Dubey, R. K., & Pournaras, E. (2023). Democracy by Design: Perspectives for Digitally Assisted, Participatory Upgrades of Society. *Journal of Computational Science*, *71*, 102061. https://doi.org/10.1016/j.jocs.2023.102061

Herington, M. J., & van de Fliert, E. (2018). Positive Deviance in Theory and Practice: A Conceptual Review. *Deviant Behavior*, *39*(5), 664-678. https://doi.org/10.1080/01639625.2017.1286194

Hernández Quiñones, A., & Flórez Herrera, J. H. (2013). *Accountability social y democracia: El caso de la Red Latinoamericana por Ciudades Justas, Democráticas y Sustentables*. Universidad de los Andes, Centro Interdisciplinario de Estudios sobre ….

Hevia, F. J. (2016). *Monitoreo Ciudadano en México. Participación ciudadana para mejorar la provisión de servicios públicos*. Centro de Investigación y Docencia Económicas.

Heywood, A. (2015). *Key Concepts in Politics and International Relations* (2 edition). Palgrave Macmillan.

Hezri, A. (2004). Sustainability indicator system and policy processes in Malaysia: A framework for utilisation and learning. *Journal of Environmental Management*, *73*(4), 357-371.

Hezri, A. A., & Dovers, S. R. (2006). Sustainability indicators, policy and governance: Issues for ecological economics. *Ecological Economics*, *60*(1), 86-99. https://doi.org/10.1016/j.ecolecon.2005.11.019

Hoffman, M. (2012). *Applying Integrated Reporting principles in the public sector*. KPMG International. https://www.kpmg.com/Global/en/topics/corporate-reporting/better-reporting/Documents/applying-ir-principles.pdf

Holden, M. (2007). Revisiting the Local Impact of Community Indicators Projects: Sustainable Seattle as Prophet in Its Own Land. *Applied Research in Quality of Life*, *1*(3-4), 253-277. https://doi.org/10.1007/s11482-007-9020-8

Holden, M. (2013). Sustainability indicator systems within urban governance: Usability analysis of sustainability indicator systems as boundary objects. *Ecological Indicators*, *32*, 89-96. https://doi.org/10.1016/j.ecolind.2013.03.007

Holman, N. (2009). Incorporating local sustainability indicators into structures of local governance: A review of the literature. *Local Environment*, *14*(4), 365-375. https://doi.org/10.1080/13549830902783043

Hood, C. (2012). Public Management by Numbers as a Performance Enhancing Drug: Two Hypotheses. *Public Administration Review*, *72*(S1), 85-92. https://doi.org/10.111/j.1540-6210.2012.02634.x

Høyland, B., Moene, K., & Willumsen, F. (2012). The tyranny of international index rankings. *Journal of Development Economics*, *97*(1), 1-14. https://doi.org/10.1016/j.jdeveco.2011.01.007

Hsiao, P. K., De Villiers, C., Horner, C., & Oosthuizen, H. (2022). A review and synthesis of contemporary sustainability accounting research and the development of a research agenda. *Accounting & Finance*, *62*(4), 4453-4483. https://doi.org/10.1111/acfi.12936

Huovila, A., Bosch, P., & Airaksinen, M. (2019). Comparative analysis of standardized indicators for Smart sustainable cities: What indicators and standards to use and when? *Cities*, *89*, 141-153. https://doi.org/10.1016/j.cities.2019.01.029

Igalla, M., Edelenbos, J., & van Meerkerk, I. (2019). Citizens in Action, What Do They Accomplish? A Systematic Literature Review of Citizen Initiatives, Their Main Characteristics, Outcomes, and Factors. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *30*(5), 1176-1194. https://doi.org/10.1007/s11266-019-00129-0

in 't Veld, R. J. (Red.). (2010). *Knowledge Democracy*. Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-642-11381-9

Ingrams, A. (2020). Administrative Reform and the Quest for Openness: A Popperian Review of Open Government. *Administration & Society*, *52*(2), 319-340. https://doi.org/10.1177/0095399719875460

Inter-American Development Bank. (2016). *Evaluation of the IDB's Emerging and Sustainable Cities Initiative*. Office of Evaluation and Oversight.

International Integrated Reporting Council. (2008). *Consultation Draft of the International Integrated Reporting Framework*. http://integratedreporting.org/wp-content/uploads/2013/12/13-12-08-THE-INTERNATIONAL-IR-FRAMEWORK-2-1.pdf

ISO. (2014). *Sustainable development of communities—Indicators for city services and quality of life* (No. ISO 37120). https://www.iso.org/obp/ui/#iso:std:iso:37120:ed-1:v1:en

Jalisco Cómo Vamos (Regisseur). (2022, november 1). *Presentación ¿Qué has hecho, Alcalde? Ciclo 2021 - 2024* [Video recording]. https://www.youtube.com/watch?v=9iC8tT2MMus

Janssen, M., & van der Voort, H. (2016). Adaptive governance: Towards a stable, accountable and responsive government. *Government Information Quarterly*, *33*(1), 1-5. https://doi.org/10.1016/j.giq.2016.02.003

Jensen, J. O., & Elle, M. (2007). Exploring the Use of Tools for Urban Sustainability in European Cities. *Indoor and Built Environment*, *16*(3), 235-247. https://doi.org/10.1177/1420326X07079341

Jerven, M. (2013). *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It*. Cornell University Press.

Joas, M., Theobald, K., McGuinness, D., Garzillo, C., & Kuhn, S. (Red.). (2014). *Informed Cities: Making Research Work for Local Sustainability*. Routledge.

Joshi, A., & Houtzager, P. P. (2012). Widgets or Watchdogs?: Conceptual explorations in social accountability. *Public Management Review*, *14*(2), 145-162. https://doi.org/10.1080/14719037.2012.657837

Joss, S. (2012). *Tomorrow's City Today: Eco-City Indicators, Standards & Frameworks. Bellagio Conference Report*. University of Westminster. http://www.westminster.ac.uk/ecocities/projects/bellagio-project/the-bellagio-report

Joss, S. (2015). *Sustainable cities: Governing for urban innovation*. Palgrave Macmillan. https://books.google.nl/books?hl=en&lr=&id=VMfkCQAAQBAJ&oi=fnd&pg=PP1&dq=joss+governing&ots=Q1yb1AYfMv&sig=b3gSy1CZLjFvObVFqyGp8gI2dbg

Joss, S., Cowley, R., de Jong, M., Müller, B., Park, B. S., Rees, W. E., Roseland, M., & Rydin, Y. (2015). *Tomorrow's s City Today. Prospects for Standardising Sustainable Urban Development*. University of Westminster. http://www.westminster.ac.uk/__data/assets/pdf_file/0020/372710/7178_TomorrowsCitiesToday.pdf

Joss, S., & Rydin, Y. (2018). Prospects of standardizing sustainable urban development. In S. Bell & S. Morse (Red.), *Routledge Handbook of*

Sustainability Indicators (pp. 364-377). Routledge. https://doi.org/10.4324/9781315561103

Joss, S., Tomozeiu, D., & Cowley, R. (2012). Eco-city indicators: Governance challenges. *WIT Transactions on Ecology and the Environment*, *155*, 109-120. https://doi.org/10.2495/SC120101

Kahalon, R., Klein, V., Ksenofontov, I., Ullrich, J., & Wright, S. C. (2022). Mentioning the Sample's Country in the Article's Title Leads to Bias in Research Evaluation. *Social Psychological and Personality Science*, *13*(2), 352-361. https://doi.org/10.1177/19485506211024036

Kampfmann, T., Bernert, P., Lang, D. J., & Drautz, S. (2024). Governance for urban sustainability through real-world experimentation – Introducing an evaluation framework for transformative research involving public actors. *Cities*, *153*, 105301. https://doi.org/10.1016/j.cities.2024.105301

Kaur, A., & Lodhia, S. K. (2019). Sustainability accounting, accountability and reporting in the public sector: An overview and suggestions for future research. *Meditari Accountancy Research*, *27*(4), 498-504.

Kitchin, R., Lauriault, T. P., & McArdle, G. (2015). Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies, Regional Science*, *2*(1), 6-28. https://doi.org/10.1080/21681376.2014.983149

Klopp, J. M., & Petretta, D. L. (2017). The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. *Cities*, *63*, 92-97. https://doi.org/10.1016/j.cities.2016.12.019

Krank, S., & Wallbaum, H. (2011). Lessons from seven sustainability indicator programs in developing countries of Asia. *Ecological Indicators*, *11*(5), 1385-1395.

Krank, S., Wallbaum, H., & Grêt-Regamey, A. (2010). Constraints to implementation of sustainability indicator systems in five Asian cities. *Local Environment*, *15*(8), 731-742. https://doi.org/10.1080/13549839.2010.509386

Krank, S., Wallbaum, H., & Grêt-Regamey, A. (2013). Perceived contribution of indicator systems to sustainable development in developing countries. *Sustainable Development*, *21*(1), 18-29. https://doi.org/10.1002/sd.496

Krantz, V., & Gustafsson, S. (2021). Localizing the sustainable development goals through an integrated approach in municipalities: Early experiences from a Swedish forerunner. *Journal of Environmental Planning and Management*, 1-19. https://doi.org/10.1080/09640568.2021.1877642

Kremer, P., Haase, A., & Haase, D. (2019). The future of urban sustainability: Smart, efficient, green or just? Introduction to the special issue. *Sustainable Cities and Society*, *51*, 101761. https://doi.org/10.1016/j.scs.2019.101761

Kroll, A. (2015). Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. *Public Performance & Management Review*, *38*(3), 459-486. https://doi.org/10.1080/15309576.2015.1006469

Kroll, A., Neshkova, M. I., & Pandey, S. K. (2019). Spillover Effects From Customer to Citizen Orientation: How Performance Management Reforms Can Foster Public Participation. *Administration & Society*, *51*(8), 1227-1253. https://doi.org/10.1177/0095399716687341

Krueger, A. B., & Stone, A. A. (2014). Progress in measuring subjective well-being. *Science*, *346*(6205), 42-43. https://doi.org/10.1126/science.1256392

Lamprinidi, S., & Kubo, N. (2008). Debate: The Global Reporting Initiative and Public Agencies. *Public Money & Management*, *28*(6), 326-329. https://doi.org/10.1111/j.1467-9302.2008.00663.x

Lange, P., Driessen, P. P. J., Sauer, A., Bornemann, B., & Burger, P. (2013). Governing Towards Sustainability—Conceptualizing Modes of Governance. *Journal of Environmental Policy & Planning*, *15*(3), 403-425. https://doi.org/10.1080/1523908X.2013.769414

Larrinaga, C., & Bebbington, J. (2021). The pre-history of sustainability reporting: A constructivist reading. *Accounting, Auditing & Accountability Journal*, *34*(9), 162-181. https://doi.org/10.1108/AAAJ-03-2017-2872

Lehtonen, M. (2015). Indicators: Tools for informing, monitoring or controlling? In A. J. Jordan & J. R. Turnpenny (Red.), *The Tools of Policy Formulation: Actors, Capacities, Venues and Effects* (pp. 76-99). Edward Elgar Publishing.

Lehtonen, M., Sébastien, L., & Bauler, T. (2016). The multiple roles of sustainability indicators in informational governance: Between intended use and unanticipated influence. *Current Opinion in Environmental Sustainability*, *18*, 1-9. https://doi.org/10.1016/j.cosust.2015.05.009

Levy, D. L., Szejnwald Brown, H., & de Jong, M. (2010). The Contested Politics of Corporate Governance: The Case of the Global Reporting Initiative. *Business & Society*, *49*(1), 88-115. https://doi.org/10.1177/0007650309345420

Lewis, J. M. (2015). The politics and consequences of performance measurement. *Policy and Society*, *34*(1), 1-12. https://doi.org/10.1016/j.polsoc.2015.03.001

Liao, L., Warner, M. E., & Homsy, G. C. (2020). When Do Plans Matter?: Tracking Changes in Local Government Sustainability Actions From 2010 to 2015. *Journal of the American Planning Association*, *86*(1), 60-74. https://doi.org/10.1080/01944363.2019.1667262

Lijphart, A. (2012). *Patterns of democracy: Government forms and performance in thirty-six countries* (Second Edition). Yale university press.

Liu, J., Dou, Y., Batistella, M., Challies, E., Connor, T., Friis, C., Millington, J. D., Parish, E., Romulo, C. L., Silva, R. F. B., Triezenberg, H., Yang, H., Zhao, Z., Zimmerer, K. S., Huettmann, F., Treglia, M. L., Basher, Z., Chung, M. G., Herzberger, A., … Sun, J. (2018). Spillover systems in a telecoupled Anthropocene: Typology, methods, and governance for global sustainability. *Current Opinion in Environmental Sustainability*, *33*, 58-69. https://doi.org/10.1016/j.cosust.2018.04.009

Lodhia, S., Jacobs, K., & Park, Y. J. (2012). Driving Public Sector Environmental Reporting: The disclosure practices of Australian Commonwealth Departments. *Public Management Review*, *14*(5), 631-647. https://doi.org/10.1080/14719037.2011.642565

Lora, E. (2016). Using Life Satisfaction Data to Identify Urban Problems, Prioritize Local Public Expenditures and Monitor the Quality of Urban Life. In M. Rojas (Red.), *Handbook of Happiness Research in Latin America* (pp. 273-294). Springer Netherlands. http://link.springer.com/10.1007/978-94-017-7203-7

Low, S. M., Abdullah, D. F., & Khatib, S. F. A. (2023). Research trend in Sustainable Development Goals reporting: A systematic literature review. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-023-30122-6

Lozano, R., & Huisingh, D. (2011). Inter-linking issues and dimensions in sustainability reporting. *Journal of Cleaner Production*, *19*(2-3), 99-107. https://doi.org/10.1016/j.jclepro.2010.01.004

LUBW. (2015). *N!-Berichte für Kommunen. Leitfaden zur Erstellung von Nachhaltigkeitsberichten in kleinen und mittleren Kommunen.* (No. 2nd edition). Ministerium für Umwelt,Klima und Energiewirtschaft Baden-Württemberg. https://www.lubw.baden-wuerttemberg.de/servlet/is/240401/n-berichte_fuer_kommunen.pdf?command=downloadContent&filename=n-berichte_fuer_kommunen.pdf

Lyytimäki, J. (2019). Thermostat or thermometer? A Finnish perspective on the overloaded role of sustainability indicators in societal transition. *Sustainable Development*, *27*(5), 817-825. https://doi.org/10.1002/sd.1941

Lyytimäki, J., & Rosenström, U. (2008). Skeletons out of the closet: Effectiveness of conceptual frameworks for communicating sustainable development indicators. *Sustainable Development*, *16*(5), 301-313. https://doi.org/10.1002/sd.330

Lyytimäki, J., Salo, H., Lepenies, R., Büttner, L., & Mustajoki, J. (2020). Risks of producing and using indicators of sustainable development goals. *Sustainable Development*, sd.2102. https://doi.org/10.1002/sd.2102

Lyytimäki, J., Tapio, P., Varho, V., & Söderman, T. (2013). The use, non-use and misuse of indicators in sustainability assessment and communication. *International Journal of Sustainable Development & World Ecology*, *20*(5), 385-393. https://doi.org/10.1080/13504509.2013.834524

Manes-Rossi, F., Brusca, I., Orelli, R. L., Lorson, P. C., & Haustein, E. (2023). Features and drivers of citizen participation: Insights from participatory budgeting in three European cities. *Public Management Review*, *25*(2), 201-223. https://doi.org/10.1080/14719037.2021.1963821

Manes-Rossi, F., Nicolò, G., & Argento, D. (2020). Non-financial reporting formats in public sector organizations: A structured literature review. *Journal of Public Budgeting, Accounting & Financial Management*, *32*(4), 639-669.

Marcuccio, M., & Steccolini, I. (2005). Social and environmental reporting in local authorities. *Public Management Review*, *7*(2), 155-176. https://doi.org/10.1080/14719030500090444

Marcuccio, M., & Steccolini, I. (2009). Patterns of voluntary extended performance reporting in Italian local authorities. *International Journal of Public Sector Management*, *22*(2), 146-167. https://doi.org/10.1108/09513550910934547

Marin, P. de L. (2016). *Sistemas de gestão para resultados no setor público: Intersecções entre política, governança e desempenho nas prefeituras de Rio de Janeiro e São Paulo* [Doctoral Thesis, Fundação Getulio Vargas]. http://bibliotecadigital.fgv.br/dspace/handle/10438/15961

Martinez, J., Mikkelsen, C. A., & Phillips, R. (2021). Introduction: Quality of Life and Sustainability, Socio-spatial, and Multidisciplinary Perspectives. In J. Martinez, C. A. Mikkelsen, & R. Phillips (Red.), *Handbook of Quality of Life and Sustainability* (pp. 1-14). Springer International Publishing. https://doi.org/10.1007/978-3-030-50540-0

McLean, B. L., & Borén, T. (2014). Barriers to implementing sustainability locally: A case study of policy immobilities. *Local Environment*, *ahead-of-print*, 1-18.

Mees, H., & Driessen, P. (2019). A framework for assessing the accountability of local governance arrangements for adaptation to climate change. *Journal of Environmental Planning and Management*, *62*(4), 671-691. https://doi.org/10.1080/09640568.2018.1428184

Meijer, A., 't Hart, P., & Worthy, B. (2015). Assessing Government Transparency An Interpretive Framework. *Administration & Society*, 0095399715598341.

Meijering, J. V., Kern, K., & Tobi, H. (2014). Identifying the methodological characteristics of European green city rankings. *Ecological Indicators*, *43*, 132-142. https://doi.org/10.1016/j.ecolind.2014.02.026

Mendonça, P. M. E., & Grandé, H. M. (2023). Civil Society in Latin America: Experiments, Resilience, New Utopias. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 900-910. https://doi.org/10.1007/s11266-023-00599-3

Merino-Saum, A., Halla, P. J., Superti, V., Boesch, A., & Binder, C. R. (2020). Indicators for urban sustainability: Key lessons from a systematic analysis of 67 measurement initiatives. *Ecological Indicators*, *119*, 106879. https://doi.org/10.1016/j.ecolind.2020.106879

Meuleman, L., & Niestroy, I. (2015). Common But Differentiated Governance: A Metagovernance Approach to Make the SDGs Work. *Sustainability*, *7*(9), 12295-12321. https://doi.org/10.3390/su70912295

Michener, G. (2015). Policy Evaluation via Composite Indexes: Qualitative Lessons from International Transparency Policy Indexes. *World Development*, *74*, 184-196. https://doi.org/10.1016/j.worlddev.2015.04.016

Milne, M. J., & Gray, R. (2013). W(h)ither Ecology? The Triple Bottom Line, the Global Reporting Initiative, and Corporate Sustainability Reporting.

Journal of Business Ethics, *118*(1), 13-29. https://doi.org/10.1007/s10551-012-1543-8

Mitchell, M., Curtis, A., & Davidson, P. (2008). Evaluating the process of triple bottom line reporting: Increasing the potential for change. *Local Environment*, *13*(2), 67-80. https://doi.org/10.1080/13549830701581937

Mol, A., Van Schie, V., & Budding, T. (2024). Drivers of sustainability reporting by local governments over time: A structured literature review. *Financial Accountability & Management*, faam.12407. https://doi.org/10.1111/faam.12407

Moldan, B., Janoušková, S., & Hák, T. (2012). How to understand and measure environmental sustainability: Indicators and targets. *Ecological Indicators*, *17*, 4-13. https://doi.org/10.1016/j.ecolind.2011.04.033

Moreno Pires, S., Fidélis, T., & Ramos, T. B. (2014). Measuring and comparing local sustainable development through common indicators: Constraints and achievements in practice. *Cities*, *39*, 1-9. https://doi.org/10.1016/j.cities.2014.02.003

Moreno Pires, S., Magee, L., & Holden, M. (2017). Learning from community indicators movements: Towards a citizen-powered urban data revolution. *Environment and Planning C: Politics and Space*, *35*(7), 1304-1337. https://doi.org/10.1177/2399654417691512

Moreno-Pires, S., & Fidélis, T. (2012). A proposal to explore the role of sustainability indicators in local governance contexts: The case of Palmela, Portugal. *Ecological Indicators*, *23*, 608-615. https://doi.org/10.1016/j.ecolind.2012.05.003

Mori, K., & Yamashita, T. (2015). Methodological framework of sustainability assessment in City Sustainability Index (CSI): A concept of constraint and maximisation indicators. *Habitat International*, *45, Part 1*, 10-14. https://doi.org/10.1016/j.habitatint.2014.06.013

Morita, K., Okitasari, M., & Masuda, H. (2020). Analysis of national and local governance systems to achieve the sustainable development goals: Case studies of Japan and Indonesia. *Sustainability Science*, *15*(1), 179-202. https://doi.org/10.1007/s11625-019-00739-z

Muñoz Ocampo, R. J., & Álvarez Méndez, J. F. (2023). *Propuesta para modificar la Ley 131 de 1994, estableciendo un marco normativo a los programas de gobierno y su método de control* [masterThesis, Universidad EAFIT]. http://repository.eafit.edu.co/handle/10784/32548

Navarro Galera, A., de los Ríos Berjillos, A., Ruiz Lozano, M., & Tirado Valencia, P. (2014). Transparency of sustainability information in local governments: English-speaking and Nordic cross-country analysis. *Journal of Cleaner Production*, *64*, 495-504. https://doi.org/10.1016/j.jclepro.2013.07.038

Navarro Galera, A., Ortiz Rodriguez, D., & López Hernández, A. M. (2008). Identifying barriers to the application of standardized performance indicators in local government. *Public Management Review*, *10*(2), 241-262. https://doi.org/10.1080/14719030801928706

Navarro-Galera, A., Ruiz-Lozano, M., Tirado-Valencia, P., & Ríos-Berjillos, A. de los. (2017). Promoting Sustainability Transparency in European Local Governments: An Empirical Analysis Based on Administrative Cultures. *Sustainability*, *9*(3), 432. https://doi.org/10.3390/su9030432

Ness, B., Urbel-Piirsalu, E., Anderberg, S., & Olsson, L. (2007). Categorising tools for sustainability assessment. *Ecological Economics*, *60*(3), 498-508.

Niemann, L., & Hoppe, T. (2018). Sustainability reporting by local governments: A magic tool? Lessons on use and usefulness from European pioneers. *Public Management Review*, *20*(1), 201-223. https://doi.org/10.1080/14719037.2017.1293149

Niemann, L., & Hoppe, T. (2021). How to Sustain Sustainability Monitoring in Cities: Lessons from 49 Community Indicator Initiatives across 10 Latin American Countries. *Sustainability*, *13*(9), 5133. https://doi.org/10.3390/su13095133

Niemann, L., Hoppe, T., & Coenen, F. (2017). On the Benefits of Using Process Indicators in Local Sustainability Monitoring: Lessons from a Dutch municipal ranking (1999–2014). *Environmental Policy and Governance*, *27*(1), 28-44. https://doi.org/10.1002/eet.1733

Nunes Silva, C. (Red.). (2020). *Contemporary Trends In Local Governance. Reform, Cooperation and Citizen Participation*. Springer. https://doi.org/10.1007/978-3-030-52516-3

OHCHR. (2015). *Proposed indicators and metadata for SDGs 10 and 16. Discussion Paper*. Office of the United Nations High Commissioner for Human Rights.

Orellana, A., & Vicuña, M. (2019). Requerimientos de indicadores y estándares urbanos para una descentralización regional efectiva. In H. von Baer & N. Bravo (Red.), *Desarrollo territorial colaborativo: Descentralizando poder, competencias y recursos* (pp. 557-570). Ediciones Universidad de La Frontera.

Ortiz-Moya, F., & Reggiani, M. (2023). Contributions of the voluntary local review process to policy integration: Evidence from frontrunner cities. *Npj Urban Sustainability*, *3*(1), 1-8. https://doi.org/10.1038/s42949-023-00101-4

Park, A. Y. S., & Krause, R. M. (2021). Exploring the landscape of sustainability performance management systems in U.S. local governments. *Journal of Environmental Management*, *279*, 111764. https://doi.org/10.1016/j.jenvman.2020.111764

Parkhurst, J. O. (2017). *The politics of evidence: From evidence-based policy to the good governance of evidence*. Routledge.

Parris, T. M., & Kates, R. W. (2003). Characterizing and measuring sustainable development. *Annual Review of environment and resources*, *28*(1), 559-586.

Perego, P., Kennedy, S., & Whiteman, G. (2016). A lot of icing but little cake? Taking integrated reporting forward. *Journal of Cleaner Production*. https://doi.org/10.1016/j.jclepro.2016.01.106

Peters, J., Langbein, J., & Roberts, G. (2018). Generalization in the Tropics – Development Policy, Randomized Controlled Trials, and External Validity. *The World Bank Research Observer*, *33*(1), 34-64. https://doi.org/10.1093/wbro/lkx005

Phillips, R. (2003). *Community Indicators*. American Planning Association.

Pintér, L., Hardi, P., Martinuzzi, A., & Hall, J. (2012). Bellagio STAMP: Principles for sustainability assessment and measurement. *Ecological Indicators*, *17*, 20-28. https://doi.org/10.1016/j.ecolind.2011.07.001

Pirannejad, A., & Ingrams, A. (2022). Open Government Maturity Models: A Global Comparison. *Social Science Computer Review*, 08944393211063107. https://doi.org/10.1177/08944393211063107

Pitts, D. W., & Fernandez, S. (2009). The State of Public Management Research: An Analysis of Scope and Methodology. *International Public Management Journal*, *12*(4), 399-420. https://doi.org/10.1080/10967490903328170

Plawitzki, J. K. (2010). *Welcher Voraussetzungen bedarf es in deutschen Kommunen für die Erstellung eines Nachhaltigkeitsberichtes? Eine explorative Untersuchung in sechs Fallstudien* [Bachelor thesis]. Leuphana Universität.

Pogrebinschi, T. (2018). Deliberative Democracy in Latin America. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, & M. Warren (Red.), *The Oxford Handbook of Deliberative Democracy* (pp. 829-841). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198747369.013.20

Pogrebinschi, T. (2021). *Thirty Years of Democratic Innovations in Latin America*. WZB Berlin Social Science Center. https://www.econstor.eu/handle/10419/235143

Pogrebinschi, T. (2023). *Innovating Democracy?: The Means and Ends of Citizen Participation in Latin America*. Cambridge University Press. https://doi.org/10.1017/9781108690010

Pogrebinschi, T., & Ross, M. (2019). Democratic innovations in Latin America. In *Handbook of Democratic Innovation and Governance* (pp. 389-403). Edward Elgar Publishing. https://www.elgaronline.com/display/edcoll/9781786433855/9781786433855.00038.xml

Pollitt, C. (2003). Public management reform: Reliable knowledge and international experience. *OECD Journal on Budgeting*, *3*(3), 121-136.

Pollitt, C. (2006). Performance Information for Democracy. The Missing Link? *Evaluation*, *12*(1), 38-55.

Pollitt, C. (2011). 'Moderation in All Things': International Comparisons of Governance Quality. *Financial Accountability & Management*, *27*(4), 437-457. https://doi.org/10.1111/j.1468-0408.2011.00532.x

Pollitt, C. (2013). The logics of performance management. *Evaluation*, *19*(4), 346-363. https://doi.org/10.1177/1356389013505040

Pollitt, C. (2018). Performance management 40 years on: A review. Some key decisions and consequences. *Public Money & Management*, *38*(3), 167-174. https://doi.org/10.1080/09540962.2017.1407129

Pollitt, C., & Bouckaert, G. (2011). *Public Management Reform: A Comparative Analysis* (Third Edition). Oxford University Press.

Pollitt, C., & Hupe, P. (2011). Talking About Government. The role of magic concepts. *Public Management Review*, *13*(5), 641-658. https://doi.org/10.1080/14719037.2010.532963

Porto de Oliveira, O. (2017). *International Policy Diffusion and Participatory Budgeting*. Springer International Publishing. https://doi.org/10.1007/978-3-319-43337-0

Pozzebon, M., & Mailhot, C. (2012). Citizens Engaged to Improve the Sustainability and Quality of Life of Their Cities: The Case of Nossa Sao Paulo. *Journal of Change Management*, *12*(3), 301-321. https://doi.org/10.1080/14697017.2012.673072

Programa Cidades Sustentáveis. (2020). *Guia para Elaboração do Plano de Metas*. https://www.cidadessustentaveis.org.br/arquivos/Publicacoes/Guia_para_Elaboracao_do_Plano_de_Metas.pdf

Pupphachai, U., & Zuidema, C. (2021). Are small municipalities prepared to use SIs? The case of Thailand. *Environmental and Sustainability Indicators*, 100162. https://doi.org/10.1016/j.indic.2021.100162

Ramos, T. B. (2019). Sustainability Assessment: Exploring the Frontiers and Paradigms of Indicator Approaches. *Sustainability*, *11*(3), Article 3. https://doi.org/10.3390/su11030824

Ramos, T. B., & Caeiro, S. (2010). Meta-performance evaluation of sustainability indicators. *Ecological Indicators*, *10*(2), 157-166. https://doi.org/10.1016/j.ecolind.2009.04.008

Reed, M., Fraser, E., & Dougill, A. (2006). An adaptive learning process for developing and applying sustainability indicators with local communities. *Ecological Economics*, *59*(4), 406-418. https://doi.org/10.1016/j.ecolecon.2005.11.008

Rees, W. E. (1997). Is 'sustainable city' an Oxymoron? *Local Environment*, *2*(3), 303-310. https://doi.org/10.1080/13549839708725535

Reporting. (2024). In *Cambridge*. Cambridge University Press. https://dictionary.cambridge.org/dictionary/english/reporting

Rey, E., Laprise, M., & Lufkin, S. (2022). *Neighbourhoods in Transition. Brownfield Regeneration in European Metropolitan Areas*. Springer International Publishing. https://doi.org/10.1007/978-3-030-82208-8_5

Rey Salamanca, F. (2015). *Voto programático y programas de gobierno en Colombia: Garantías para su cumplimiento* (Primera edición). Editorial Universidad del Rosario.

Rimmel, G., Aras, G., Baboukardos, D., Krasodomska, J., Nielsen, C., & Schiemann, F. (2024a). Multifaceted, complex, and multilayered aspects of sustainability reporting research: An introduction. In G. Rimmel, G. Aras, D. Baboukardos, J. Krasodomska, C. Nielsen, & F. Schiemann (Red.), *Research Handbook on Sustainability Reporting* (pp. 2-11). Edward Elgar Publishing. https://doi.org/10.4337/9781035316267.00008

Rimmel, G., Aras, G., Baboukardos, D., Krasodomska, J., Nielsen, C., & Schiemann, F. (Red.). (2024b). *Research Handbook on Sustainability Reporting*. Edward Elgar Publishing.

Rinne, J., Lyytimäki, J., & Kautto, P. (2012). From sustainability to well-being: Lessons learned from the use of sustainable development indicators at national and EU level. *Ecological Indicators, 35,* 35-42.

Romanutti, V. (2012a). *Instrumentos de rendición de cuentas y participación ciudadana. Aprendizajes en América Latina*. Red Ciudadana Nuestra Córdoba. http://www.nuestracordoba.org.ar/sites/default/files/Plan_Metas_Aprendizajes_en_America_Latina.pdf

Romanutti, V. (2012b). *Plan de Metas en Córdoba. Una experiencia de incidencia colectiva*. Red Ciudadana Nuestra Córdoba.

Romanutti, V., & Cáceres, P. (2020). Innovación institucional y transformación democrática. Ocho años de implementación del plan de metas de gobierno en la ciudad de Córdoba. *Administración Pública y Sociedad*, *10*(Julio-Diciembre), 252-264.

Roseland, M., Stout, M., & Spiliotopoulou, M. (2024). The Global Context for Sustainability. In *Toward Sustainable Communities. Solutions for Citizens and Their Governments* (Fifth Edition, pp. 3-22). New Society Publishers. https://newsociety.com/products/9780865719743

Rosvall, M., Gustafsson, M., & Åberg, M. (2023). Strategic visions for local sustainability transition: Measuring maturity in Swedish municipalities. *Journal of Environmental Policy & Planning*, *25*(5), 539-553. https://doi.org/10.1080/1523908X.2023.2218273

Runesson, E., & Samani, N. (2024). Reporting for sustainability. A presentation of challenges. In M. Bruncevic, S. Fellman, A. Styhre, & M. Söderbom (Red.), *Towards a sustainable world: Academic insights and perspectives* (pp. 141-158). University of Gothenburg.

Sáez, L. (2012). *Methods in Governance Research: A Review of Research Approaches*. (Working Paper No. 17). Effective States and Inclusive Development Research Centre (ESID), The University of Manchester. http://r4d.dfid.gov.uk/PDF/Outputs/ESID/esid_wp_17_saez.pdf

Schneider, C. Q., & Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge University Press.

Schraven, D., Joss, S., & de Jong, M. (2021). Past, present, future: Engagement with sustainable urban development through 35 city labels in the scientific literature 1990–2019. *Journal of Cleaner Production*, *292*, 125924. https://doi.org/10.1016/j.jclepro.2021.125924

Scrivens, K., & Iasiello, B. (2010). *Indicators of "societal progress". Lessons from international experiences*. OECD Publishing.

SDSN. (2019). *Counting on the World to Act. A Roadmap for Governments to Achieve Modern Data Systems for Sustainable Development*. Sustainable Development Solutions Network.

Seaford, C. (2013). The Multiple Uses of Subjective Well-Being Indicators. *Social Indicators Research*, *114*(1), 29-43. https://doi.org/10.1007/s11205-013-0381-0

Sébastien, L., & Bauler, T. (2013). Use and influence of composite indicators for sustainable development at the EU-level. *Ecological Indicators*, *35*, 3-12.

Sharifi, A. (2016). A critical review of selected tools for assessing community resilience. *Ecological Indicators*, *69*, 629-647. https://doi.org/10.1016/j.ecolind.2016.05.023

Sharifi, A. (2019). A critical review of selected smart city assessment tools and indicator sets. *Journal of Cleaner Production*, *233*, 1269-1283. https://doi.org/10.1016/j.jclepro.2019.06.172

Sharifi, A. (2020). A typology of smart city assessment tools and indicator sets. *Sustainable Cities and Society*, *53*, 101936. https://doi.org/10.1016/j.scs.2019.101936

Sharifi, A., Allam, Z., Bibri, S. E., & Khavarian-Garmsir, A. R. (2024). Smart cities and sustainable development goals (SDGs): A systematic literature review of co-benefits and trade-offs. *Cities*, *146*, 104659. https://doi.org/10.1016/j.cities.2023.104659

Sheehy, B., & Farneti, F. (2021). Corporate social responsibility, sustainability, sustainable development and corporate sustainability: What is the difference, and does it matter? *Sustainability*, *13*(11), 5965. https://doi.org/10.3390/su13115965

Shen, L.-Y., Jorge Ochoa, J., Shah, M. N., & Zhang, X. (2011). The application of urban sustainability indicators – A comparison between various practices. *Habitat International*, *35*(1), 17-29. https://doi.org/10.1016/j.habitatint.2010.03.006

Sirgy, M. J. (2022). *Community Quality-of-Life Indicators: A Guide for Community Indicators Projects*. Springer International Publishing. https://doi.org/10.1007/978-3-031-10208-0

Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, *18*(2-3), 277-310. https://doi.org/10.1080/01900699508525011

Soto-González, E. (2020). *Contribución de Jalisco Cómo Vamos a la gobernanza. Aporte público y lecciones aprendidas a partir del trabajo con el Gobierno de Jalisco 2013-2018.* [Trabajo de maestria, Instituto Tecnológico y de Estudios Superiores de Occidente]. http://rei.iteso.mx/handle/11117/6218

Spada, P., & Paulson, L. (2023). Measuring the effect of collective intelligence processes that leverage participation and deliberation. In S. Boucher, C. A. Hallin, & L. Paulson (Red.), *The Routledge Handbook of Collective Intelligence for Democracy and Governance* (pp. 78-107). Routledge. https://doi.org/10.4324/9781003215929

Spangenberg, J. H. (2002). Institutional sustainability indicators: An analysis of the institutions in Agenda 21 and a draft set of indicators for monitoring their effectivity. *Sustainable Development*, *10*(2), 103-115. https://doi.org/10.1002/sd.184

Springett, D., & Redclift, M. (2017). History and evolution of the concept. In M. Redclift & D. Springett (Red.), *Routledge International Handbook of Sustainable Development* (pp. 3-38). Routledge.

Stacchezzini, R., Melloni, G., & Lai, A. (2016). Sustainability management and reporting: The role of integrated reporting for communicating corporate sustainability management. *Journal of Cleaner Production*. https://doi.org/10.1016/j.jclepro.2016.01.109

Stadt Freiburg. (2014). *1. Freiburger Nachhaltigkeitsbericht. Beispielhafter Ausschnitt zur Darstellung des Nachhaltigkeitsprozesses*.

Steccolini, I. (2004). Is the annual report an accountability medium? An empirical investigation into Italian local governments. *Financial Accountability & Management*, *20*(3), 327-350.

Stockmann, R., Meyer, W., & Szentmarjay, L. (Red.). (2022). *The Institutionalisation of Evaluation in the Americas*. Springer International Publishing. https://doi.org/10.1007/978-3-030-81139-6

Stubbs, W., & Higgins, C. (2014). Integrated Reporting and internal mechanisms of change. *Accounting, Auditing & Accountability Journal*, *27*(7), 1068-1089. https://doi.org/10.1108/AAAJ-03-2013-1279

Sustainable Development Solutions Network. (2015). *Indicators and a Monitoring Framework for the Sustainable Development Goals. Launching a data revolution for the SDGs*. Sustainable Development Solutions Network. http://unsdsn.org/wp-content/uploads/2015/05/150612-FINAL-SDSN-Indicator-Report1.pdf

Swiss Federal Office for Spatial Development. (2005). *Quality Criteria for a Sustainable Development Process. Guidance for Sustainable Development Actors within Public Communities*. http://www.are.admin.ch/dokumentation/publikationen/00014/index.html?lang=en

Tanguay, G. A., Rajaonson, J., & Bilodeau, P.-K. (2023). Linking quality of life and sustainability in Canadian cities. *Cities*, *143*, 104608. https://doi.org/10.1016/j.cities.2023.104608

Tanguay, G. A., Rajaonson, J., Lefebvre, J.-F., & Lanoie, P. (2010). Measuring the sustainability of cities: An analysis of the use of local indicators. *Ecological Indicators*, *10*(2), 407-418.

Thomas, R., Hsu, A., & Weinfurter, A. (2020). Sustainable and inclusive – Evaluating urban sustainability indicators' suitability for measuring progress towards SDG-11. *Environment and Planning B: Urban Analytics and City Science*, 239980832097540. https://doi.org/10.1177/2399808320975404

Tillema, S. (2010). Public sector benchmarking and performance improvement: What is the link and can it be improved? *Public Money & Management*, *30*(1), 69-75. https://doi.org/10.1080/09540960903492414

Tozer, L., Hörschelmann, K., Anguelovski, I., Bulkeley, H., & Lazova, Y. (2020). Whose city? Whose nature? Towards inclusive nature-based solution governance. *Cities*, *107*, 102892. https://doi.org/10.1016/j.cities.2020.102892

Tsoukas, H. (1997). The tyranny of light: The temptations and the paradoxes of the information society. *Futures*, *29*(9), 827-843.

Turnhout, E., Hisschemöller, M., & Eijsackers, H. (2007). Ecological indicators: Between the two fires of science and policy. *Ecological indicators*, *7*(2), 215-228.

United Cities and Local Governments. (2014). *How to localize targets and indicators of the post-2015 agenda? UCLG inputs for the Global Taskforce Working Paper*. UCLG. http://www.uclg.org/sites/default/files/localization_targets_indicator_web.pdf

United Nations. (2012, april 23). *'Our Struggle for Global Sustainability Will Be Won or Lost in Cities,' Says Secretary-General*. http://www.un.org/press/en/2012/sgsm14249.doc.htm

United States Government Accountability Office. (2011). *Key Indicator Systems: Experiences of Other National and Sub-National Systems Offer Insights for the United States* (No. March). DIANE Publishing. http://www.gao.gov/products/GAO-11-396

Uribe Mendoza, C. (2016). La activación de la revocatoria de mandato en el ámbito municipal en Colombia. Lecciones del caso de Bogotá. *Estudios Políticos*, *48*, 179-200.

Valencia, S. C., Simon, D., Croese, S., Nordqvist, J., Oloko, M., Sharma, T., Taylor Buck, N., & Versace, I. (2019). Adapting the Sustainable Development Goals and the New Urban Agenda to the city level: Initial reflections from a comparative research project. *International Journal of Urban Sustainable Development*, *11*(1), 4-23. https://doi.org/10.1080/19463138.2019.1573172

van Acker, W., & Bouckaert, G. (2018). What makes public sector innovations survive? An exploratory study of the influence of feedback, accountability and learning. *International Review of Administrative Sciences*, *84*(2), 249-268. https://doi.org/10.1177/0020852317700481

Van de Walle, S., & Cornelissen, F. (2014). Performance Reporting. In M. Bovens, R. Goodin, & T. Schillemans (Red.), *The Oxford Handbook of Public Accountability* (pp. 441-455). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641253.013.0009

van der Heijden, J. (2014). *Governance for Urban Sustainability and Resilience*. Edward Elgar Publishing. https://doi.org/10.4337/9781782548133

Van Dooren, W., Bouckaert, G., & Halligan, J. (2015). *Performance management in the public sector* (Second edition). Routledge.

Velte, P. (2022). Meta-analyses on Corporate Social Responsibility (CSR): A literature review. *Management Review Quarterly*, *72*(3), 627-675. https://doi.org/10.1007/s11301-021-00211-2

Vermeulen, W. (2018). Substantiating the rough consensus on concept of sustainable development as point of departure for indicator development. In S. Bell & S. Morse (Red.), *Routledge Handbook of Sustainability Indicators* (pp. 59-90). Routledge. https://doi.org/10.4324/9781315561103

Villela, M., & Souza, R. (2021). Programas de metas municipais sob olhar do Tribunal de Justiça de São Paulo. *Revista Digital de Direito*

Administrativo, *8*(1), Article 1. https://doi.org/10.11606/issn.2319-0558.v8i1p205-226

Waas, T., Hugé, J., Block, T., Wright, T., Benitez-Capistros, F., & Verbruggen, A. (2014). Sustainability assessment and indicators: Tools in a decision-making strategy for sustainable development. *Sustainability*, *6*(9), 5512-5534.

Walker, W. E., Rahman, S. A., & Cave, J. (2001). Adaptive policies, policy analysis, and policy-making. *European Journal of Operational Research*, *128*(2), 282-289. https://doi.org/10.1016/S0377-2217(00)00071-0

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, *39*(5), 426-431.

Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence how evaluations affect DARE. *American Journal of Evaluation*, *26*(1), 12-30.

Wiedmann, T., & Allen, C. (2021). City footprints and SDGs provide untapped potential for assessing city sustainability. *Nature Communications*, *12*(1), 3758. https://doi.org/10.1038/s41467-021-23968-2

Willems, T., & Van Dooren, W. (2012). Coming to Terms with Accountability. *Public Management Review*, *14*(7), 1011-1036. https://doi.org/10.1080/14719037.2012.662446

Williams, B., Wilmshurst, T., & Clift, R. (2011). Sustainability reporting by local government in Australia: Current and future prospects. *Accounting Forum*, *35*, 176-186.

Wilson, J. Q. (1991). *Bureaucracy: What government agencies do and why they do it*. Hachette UK.

Wittmayer, J. M., van Steenbergen, F., Rok, A., & Roorda, C. (2015). Governing sustainability: A dialogue between Local Agenda 21 and transition management. *Local Environment*, 1-17. https://doi.org/10.1080/13549839.2015.1050658

Wood, D. M. (2016). Community indicators and collective impact: Facilitating change. *Community Development*, *47*(2), 194-208. https://doi.org/10.1080/15575330.2015.1133685

World Bank. (2017). *World Development Report 2017: Governance and the Law*. The World Bank. https://doi.org/10.1596/978-1-4648-0950-7

World Commission on Environment and Development. (1987). *Our Common Future*. Oxford University Press.

Wray, L., Stevens, C., & Holden, M. (2017). The History, Status and Future of the Community Indicators Movement. In M. Holden, R. Phillips, & C. Stevens (Red.), *Community Quality-of-Life Indicators: Best Cases VII* (pp. 1-16). Springer.

Wu, X., Fu, B., Wang, S., Song, S., Lusseau, D., Liu, Y., Xu, Z., & Liu, J. (2023). Bleak prospects and targeted actions for achieving the Sustainable Development Goals. *Science Bulletin*, *68*(22), 2838-2848. https://doi.org/10.1016/j.scib.2023.09.010

Yang, K., & Hsieh, J. Y. (2007). Managerial Effectiveness of Government Performance Measurement: Testing a Middle-Range Model. *Public Administration Review*, *67*(5), 861-879. https://doi.org/10.1111/j.1540-6210.2007.00774.x

Yigitcanlar, T., & Lönnqvist, A. (2013). Benchmarking knowledge-based urban development performance: Results from the international comparison of Helsinki. *Cities*, *31*, 357-369.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (Sixth edition). SAGE.

Zeemering, E. S. (2018). Sustainability management, strategy and reform in local government. *Public Management Review*, *20*(1), 136-153. https://doi.org/10.1080/14719037.2017.1293148

Zeng, Y., Maxwell, S., Runting, R. K., Venter, O., Watson, J. E. M., & Carrasco, L. R. (2020). Environmental destruction not avoided with the Sustainable Development Goals. *Nature Sustainability*, *3*(10), Article 10. https://doi.org/10.1038/s41893-020-0555-0

Zhou, N., & Williams, C. (2013). *An International Review of Eco-City Theory, Indicators, and Case Studies*. Ernest Orlando Lawrence Berkeley National Laboratory. http://eetd.lbl.gov/sites/all/files/lc_eco-cities.pdf

Zinkernagel, R., Evans, J., & Neij, L. (2018). Applying the SDGs to Cities: Business as Usual or a New Dawn? *Sustainability*, *10*(9), 3201. https://doi.org/10.3390/su10093201

8. List of publications

The four empirical studies underpinning this dissertation have each been published as academic article in a peer-reviewed journal. Table 8.1 contains publication details of each article including co-authorships. The references of all articles have been integrated into one integrated reference list. For some articles, the online repositories of the respective journals also contain supplementary material which is not included in this dissertation.

Table 8.1 *Details of academic articles*

Chapter	Details of academic article	
2	Research question:	How have sustainability reporting practices evolved in pioneering European city governments, and what are their effects?
	Article title:	**Sustainability reporting by local governments: a magic tool? Lessons on use and usefulness from European pioneers**
	Journal:	Public Management Review
	Year:	2018
	DOI:	http://dx.doi.org/10.1080/14719037.2017.1293149
	Authors:	Ludger Niemann & Thomas Hoppe
3	Research question:	How have sustainability reporting practices by community indicator initiatives evolved in Latin American cities, and what are their effects?
	Article title:	**How to Sustain Sustainability Monitoring in Cities: Lessons from 49 Community Indicator Initiatives across 10 Latin American Countries**
	Journal:	Sustainability
	Year:	2021
	DOI:	https://www.mdpi.com/2071-1050/13/9/5133
	Supplementary material	https://www.mdpi.com/article/10.3390/su13095133/s1
	Authors:	Ludger Niemann & Thomas Hoppe

4	Research question:	What are the strengths and limitations of reporting initiatives addressing the sustainability policies of European local governments via voluntary benchmarking?
	Article title:	**On the Benefits of Using Process Indicators in Local Sustainability Monitoring: Lessons from a Dutch municipal ranking (1999–2014)**
	Journal:	Environmental Policy & Governance
	Year:	2017
	DOI:	http://onlinelibrary.wiley.com/doi/10.1002/eet.1733/epdf
	Authors:	Ludger Niemann, Thomas Hoppe & Frans Coenen
5	Research question:	What are the strengths and limitations of reporting initiatives addressing the sustainability policies of Latin American local governments via accountability obligations?
	Article title:	**Should mayors be accountable for election promises? Effects of compulsory goal setting and reporting requirements on sustainability governance in four Latin American cities**
	Journal:	Frontiers in Sustainable Cities
	Year:	2025
	DOI:	https://www.frontiersin.org/journals/sustainable-cities/articles/10.3389/frsc.2025.1450933
	Supplementary material	https://www.frontiersin.org/articles/10.3389/frsc.2025.1450933/full#supplementary-material
	Authors:	Ludger Niemann & Thomas Hoppe

Other relevant publications:

- Niemann, L., & Hoppe, T. (2022). Monitoreo local de la sostenibilidad: Iniciativas y aportes de la sociedad civil en ciudades latinoamericanas. In D. P. Petri (Ed.), *Negociación Internacional en América Latina* (pp. 145-168). FLACSO Costa Rica. https://www.flacso.ac.cr/es/publicaciones/libros/218-negociacion-internacional-en-america-latina
- Niemann, L., Morssinkhof, S., van der Linden, M. J., & de Vries, K. (2025). Data Requests in Value Chains: The Effects of Corporate Sustainability Reporting on SMEs in The Netherlands. *Sustainability*, *17*(17), 8029. https://doi.org/10.3390/su17178029

9. Short curriculum vitae

After growing up in Germany and spending a year at school in Mexico, Ludger Niemann began his academic journey in anthropology and archaeology before earning a Bachelor of Arts in Natural Sciences (Experimental Psychology) from the University of Cambridge in 1998. In 2002 he completed a master's degree (*Diplom-Psychologe*) at the University of Freiburg. His thesis research—a statistical meta-evaluation of mindfulness-based stress reduction courses—formed the basis for a subsequent journal article co-authored with his supervisors, which has since been cited over 7,500 times. From 2003 to 2014, he worked with NGOs such as CARE and Oxfam, as well as with the parastatal organisation GIZ. He focused on emergency relief and international development, with multi-year assignments in Afghanistan and Ecuador. In 2015, he pursued a Master's in Public Administration from Erasmus University Rotterdam. Since 2016, he has served as a lecturer, team leader, and researcher at The Hague University of Applied Sciences, working in the domains of public management, safety and security, and sustainability reporting.

Sustainability reporting is widely promoted as a pathway to better urban governance—but does more data really lead to better decisions, learning, or accountability? While indicator frameworks have proliferated globally, research remains dominated by Western contexts. This leaves the long-term dynamics in the Global South largely unexplored and casts doubt on the explanatory power of existing theories.

This dissertation bridges that gap. Through a longitudinal, comparative analysis of cities in Europe and Latin America, it investigates how reporting initiatives function within their specific ecological niches. Findings show that while reporting can support transparency and policy change, it risks becoming a bureaucratic ritual that loses traction when misaligned with local political realities.

By examining how design choices and stakeholder engagement shape real-world effects, this study moves beyond the "more data is better" narrative. It offers a framework for designing reporting initiatives that are credible, context-sensitive, and capable of producing lasting public value.

Ludger Niemann is a lecturer and researcher at The Hague University of Applied Sciences, specialising in sustainability reporting and urban governance. His work combines academic research with practical experience in NGOs and policy advice.

ISBN:
978-94-6518-221-6