Sara Takamatsu

# Bridging Global Goals to Local Realities: Exploring the Local Indicators and its Reporting of the Sustainable Development Goals in the European Union

**Master Thesis**

at the Chair for Information Systems and Information Management
(Westfälische Wilhelms-Universität, Münster)

Supervisor: Prof. Steven Van de Walle

Presented by: Sara Takamatsu

Date of Submission: 2023-06-06

# Content

# Figures

# Tables

## Abbreviations

| | |
|---|---|
| CSR | Corporate Social Responsibility |
| EU | European Union |
| GDP | Gross-domestic-product |
| HLPF | High-Level Political Forum |
| LRG | Local and Regional Governments |
| M&E | Monitoring and Evaluating |
| PI | Performance Information |
| SDG | Sustainable Development Goals |
| SR | Sustainability Reports |
| UN | United Nations |
| VLR | Voluntary Local Review |
| VNR | Voluntary National Review |

# 1 Introduction

The use of indicators as tools of performance management to assess public institutions and their actions has long been the mainstream (Heinrich, 2002; Hood, 2007; Micheli & Neely, 2010). Indicators have continued to gain momentum due to fundamental public management styles such as New Public Management (NPM) that values result-orientation, efficiency, accountability, and competition, as well as new methods such as digitalization that enables easier data collection (Bird et al., 2005; Hood, 1991; Van Der Knaap, 2006; Van Dooren et al., 2015). Using indicators is expected to increase the accountability, efficiency, and effectiveness of organizations (Cunningham & Harris, 2005). Now, concepts such as "evidence-based" and "results-based" policymaking are ubiquitous (J. Heinrich, 2007; Van Der Knaap, 2006). Some say that the "evaluation society" is fueling a "measurement-fever" where indicators and their reporting have taken center stage (Bandola-Gill & Smith, 2022; Bexell & Jönsson, 2019).

Indicators in the field of performance management are defined to be measurements of pre-defined objects collected as pieces of data and transformed into information through analysis for use in policies and programs (Van Dooren et al., 2015). On the one hand, indicators can serve a wide range of objectives each shaped uniquely to the organization in question, thus it is said that countries are keen on having indicators because of the "multiplicity of uses" (Mäkinen et al., 2018). On the other hand, this has led to an overabundance of performance data but a lack of "theoretical and empirical justifications for performance-reporting requirements" (Moynihan, 2008) as well as unclear conclusions about whether the use of indicators in performance management leads to people's expectations of improved accountability, efficiency or effectiveness (Cunningham & Harris, 2005). Despite these questions, one area within the public sector and policymaking that have been dominated by indicators is the monitoring of sustainability (Lyytimäki et al., 2013; Mair et al., 2018; Ramos & Caeiro, 2010; Scerri & James, 2010).

Sustainability has been at the forefront of every level of society and sector, from the international level to local municipalities. Defined as "meet(ing) the needs of the present without compromising the ability of future generations to meet their own needs" (United Nations, n.a.-b), the United Nations (UN) proposed "the 2030 Agenda for Sustainable Development Goals" (SDGs) seen in Figure 1, as an evidence-based framework aimed to ensure sustainability by 2030 as outlined by 17 goals and 169 targets (UNDP, 2023). These goals and targets are overseen by 231 unique indicators, otherwise known as the "global indicators" that were created over a span of 18 months, for countries to use as the "backbone of monitoring progress" (Leadership Council Sustainable Development

Solutions Network, 2015; United Nations, 2017b, n.a.-d). Although the use of these indicators is voluntary, they are said to have widespread influence from streamlining national and local policies, treaties, and projects to helping form the foundation for SDG-based investment frameworks (European Union, 2022a; ITU & DIAL, 2019). This prevalence of indicators can be understood as reflecting the importance of indicators and their centrality as performance tools for ensuring the SDGs (Mair et al., 2018).


**Figure 1 The 17 Sustainable Development Goals (United Nations, n.a.-c)**

Cooperation at all levels is necessary to ensure sustainability, but it can be said that there are exceptionally high expectations of local and regional governments (LRG) – or cities – to contribute, as they house half the global population and often are the closest in proximity to citizens (ICLEI, 2015). Not only is one of the 17 goals the creation of sustainable cities (United Nations, n.a.-d), but more than half of the SDGs cannot be achieved without the cities' cooperation (Cities Alliance, 2015). More and more LRGs are utilizing the SDG framework to develop policies and strategies, but the global indicators meant for country-level use can overgeneralize trends and endanger LRGs with wrong assumptions that could lead to ineffective policies and programs (OECD, 2020). Therefore, there is a need for local indicators in order to assess local SDG implementation appropriately and accurately that in turn can fuel better policymaking for the LRG, the countries, and the world. This then raises the question: how should the LRGs create these local indicators for such a global problem?

International organizations such as the OECD and the Joint Research Centre have created localized SDG indicator frameworks and guidelines aimed to help contextualize the

global indicator at sub-national and local levels (OECD, 2020). However, just like how there is an untrue assumption that the existence of indicators translates to the automatic improvement of whatever it is measuring (Van de Walle & Van Dooren, 2009), it can be assumed that the supply of these localized frameworks and guidelines do not lead to perfectly localized indicators although this assumption will not be tested within the scope of this research. Fundamentally, the LRGs have the great task of creating indicators that are appropriate as 'good' indicators and personalized to their local context while aligning with global indicators, targets, and goals.

On the flip side of this daunting task, this localization of the SDGs is gaining widespread support in the European Union (EU), as they have recently committed to a €500 million budget framework partnership dedicated to local governments' SDG implementation (European Union, 2022b). Nonetheless, despite all of these finances, resources, guidelines, and recommendations, current SDG progress is in "grave jeopardy" (United Nations, 2022). The realization of sustainability is a wicked problem with no clear answer, and it may be worth observing the indicator-based status quo that must answer to varying levels of authority, context, and priorities.

One way to assess the current situation is to examine periodic reviews and reporting as it is one of the most common ways of informing the public (Saner et al., 2020). In this case, there is the UN-recommended Voluntary Local Reviews (VLRs), a periodic report self-published by LRGs around the world that is "both a reporting tool to assess, monitor and present local achievements in implementing the SDGs, as well as a process to enhance and expand the political and social commitment of a variety of stakeholders to the SDGs, orienting local priorities and development planning" (United Cities and Local Governments & UN-Habitat, 2021). In essence, these VLRs are public reports that should be used for communicating, knowledge-sharing, decision-making, and ensuring SDG implementation is on track. Guidelines on how to create these the VLRs are available too, explaining SDG-based guiding principles to the specific structure of the report and this is in addition to general sustainability-related reporting guidelines for both public and private sector organizations created by international non-profit organizations that work with prominent institutions such as the European Commission (UNDESA, 2020; United Nations, n.a.-e). Indicators are, of course, an integral part of this VLR guideline and report. The UN-Habitat guideline acknowledges the need for supplementary regional indicators but does not specify how to create nor localize the indicators (United Cities and Local Governments & UN-Habitat, 2020). Therefore, there currently exists numerous VLRs with varying degrees and depth of reporting, and numerous progress-tracking methods with differing indicators thus making comparisons difficult (Schmidt-Traub et al., 2017).

## 1.1 Research Gap

This research has identified several research gaps concerning indicators and reports of local SDGs. First, there is less research conducted on critically questioning the nature and characteristics of the indicators and public reports at the LRG level within the EU. There is research focused on analyzing SDG implementation through the examination of the indicators on the VLRs and assessing the gaps between proposed indicators and guideline frameworks (Ciambra et al., 2021). In addition, indicator-related research within the EU context is specific to certain industries, a specific SDG goal, comparing national strategies, or a more general supranational level while LRG-level research has been conducted outside of Europe such as in Australia, India, and Singapore (e.g. Allen et al., 2019; ElMassah & Mohieldin, 2019; Greiling et al., n.d.; Kuc-Czarnecka et al., 2023, 2023; Leavesley et al., 2022; Steurer & Hametner, 2013; Tiwari et al., 2021). Pre-existing research shows that global indicators may oversimplify a problem that does not even have a consistent definition because of the subjectivity of the SDGs and their indicators. The research recommended that the indicators be used to hold governments to account instead of being the basis for policy creation (Mair et al., 2018). This thesis hopes to fill in this gap by assessing the nature and characteristics of the LRG's indicators within the scope of the EU.

Second, there is a lack of research examining the reports that these indicators are communicated through, namely the VLRs. In general, reports are crafted using a mix of texts and visual representations, and the representation of information can be manipulated to manage the readers' impressions (Samkin & Schneider, 2010). Additionally, reports can cause questions related to their "completeness, transparency, veracity and usefulness of the data" especially in the context of sustainability (Pennington & More, 2010). Considering pre-existing research that states that public annual reports by governments can have problems with reliability, timeliness, accessibility, and adequacy, skepticism about the authenticity of the organization is warranted as well as if reporting on sustainability actually leads to sustainability (Adams & Evans, 2004; Dando & Swift, 2003; Keerasuntonpong et al., 2019; Owen et al., 2000). Current research of the VLR as a report focuses on aspects like the structure (Ortiz-Moya et al., 2020). Therefore, this research also hopes to critically analyze the communication of the reports and how it may be currently crafting impressions.

Lastly, most of the research in this field is populated by gray literature. Numerous reports and recommendations by international organizations advise how cities should interpret the SDGs for local policymaking and indicator-creation by compiling examples, or for region-specific descriptive guidelines on how to conduct VLRs (Congress of Local and

Regional Authorities, n.a.; de Losada, 2021; Fox & Macleod, 2019; United Cities and Local Governments, 2022). The number of reports may indicate the struggles that cities uniquely face in implementing the SDGs. Therefore, this research aims to contribute to overall academic research on SDG localization in the European context.

## 1.2 Research Question

Following these research gaps, the main research question covers both the quality assessment of the indicator and the reports by asking: **What is the nature of the indicators communicated within the Voluntary Local Reviews by the local and regional governments of the European Union?**

The key subjects presented in the research question were two-fold: the nature of indicators, and the presentation of indicators in the reports. This leads to the first sub-question: *What are the characteristics of the indicators on the EU VLRs?* In answering this sub-question, an understanding of indicators as a tool of performance management was created including its qualities, advantages, and disadvantages.

In addition, a second sub-question is raised to understand the second subject of the main research question: *What are the characteristics of the VLRs as public reports?* The VLRs are, in essence, a public report as it is published by LRGs. However, it is already known that reports can be a tool for potential manipulation. Therefore, the concept of public reports including its core values as well as standards that guide report creation was explored.

## 1.3 Research Aim and Objectives

As more and more LRGs are getting involved in the VLR creation process, the aim of this research is to investigate how the LRGs are monitoring their SDG implementation by examining the indicators in their VLRs. The researcher intends to explore the core concepts of performance, indicators, and public reports that provide a foundation for the assessment of the VLRs. This research will *not* evaluate the progress or performance of the SDGs through indicators, but instead, focus on the indicators and its presentation.

To achieve this aim and answer the research and sub-questions, the objectives of this thesis are as follows:

1. Describe the core concepts and theories related to performance, indicators, and public reports to create a conceptual framework;

2. Identify and evaluate the characteristics of indicators and public reports within the VLRs through the conceptual framework;

3. Provide the results of point 2.

## 1.4 Research Structure

The remainder of this thesis is structured as followed. In Section 2, a literature review encompassing performance, indicators, and reports was conducted to contextualize these concepts. Through these identifications, a conceptual framework that framed this research was created and the details of this framework which was based on the performance measurement process were introduced in Section 3. Next, the introduction of the Voluntary Local Reviews as explained by international organizations such as the United Nations along with pre-existing analysis of the VLRs was conducted to understand what kind of research has already been conducted regarding this topic as seen in Section 4. In Section 5, the research methodology of an inductive qualitative case study was introduced, along with the coding scheme, data analysis, and the limitations of this method. In Section 6, the preliminary results of this research broken down into the six parts of what to measure, indicator selection, data collection, analysis, reports, and quality assurance were laid out, giving way to the discussion in Section 7 where relevance between the results and past research was examined. Finally, in Section 8, the answer to the research questions, as well as the implications, limitations, and significance of this thesis was discussed along with potential future research avenues.

# 2 Literature Review

In this section of the literature review, the focus will be on the conceptualization of indicators through the lens of performance management and reports. These topics can be examined through different perspectives such as statistical, political, social, or financial. To align the topic with the scope of this research, mainly the political and social perspectives will be considered. However, it is noted that the statistical and financial perspectives are equally important, as shown in pre-existing research.

First, the concept of performance will be understood within public administration. Much like indicators, performance is understood to be subjective, thus the surrounding context including public values, and principal-agent and principal-steward theories that ultimately influence the actions and motivations of actors will be discussed. Then, the different characteristics of indicators, and how they are used will be discussed. Finally, the characteristics of public reports will be discussed, including the different functions and roles they play as one of the main communication media of public institutions, as well as the methods and types of reports, predominant reporting guidelines, and how reports can be strategically utilized as a method of impression management.

In gaining a thorough understanding of these three concepts and how they relate to one another, a conceptual framework to examine the VLRs and their indicators can be created, which will be expanded upon in Section 3.

## 2.1 Conceptualizing Performance in the Public Sector

In this section, performance as analyzed in the public sector will be deconstructed as the process of performance combined with the lens of public value theory and other core values surrounding performance. Additionally, the agency stewardship theory and its influences on performance will be explained as this research will focus on LRGs that may be viewed as agents or stewards of its principal, their national government.

### 2.1.1 Process and Components of Performance

Performance in the public sector is a heavily scrutinized research field, from its roots in the early 20$^{th}$ century further surged by the introduction of New Public Management (NPM) in the late 21$^{st}$ century. Through decades of research, it is said that there are four types of performance: *production* with disregard to results, *capacity/competency* with disregard to results, *results* with disregard to production, and *sustainable results* that are both production- and result-oriented (Van Dooren et al., 2015, p. 4). From this definition,

it is evident that performance spans from the internal processes and resources of an organization to its external effects.

Performance was once considered to be a simple linear process starting with these internal processes that lead to external effects. However, it is now established that performance is much more complex and non-linear, as shown in Figure 2. “Performance as a process” is an established way of understanding the cyclical concept in which socio-economic situations birth a need for public policies, programs, or organizations that create outcomes, with differing influences at every step (Van Dooren et al., 2015, p. 20). Ultimately, the objective of performance is three-fold: to learn or for research purposes, to steer and control or for managerial purposes, or to give an account or for democratic purposes (Bird et al., 2005; Van Dooren et al., 2015). Based on the use of performance results, the emphasis on the process as well as methods will differ.

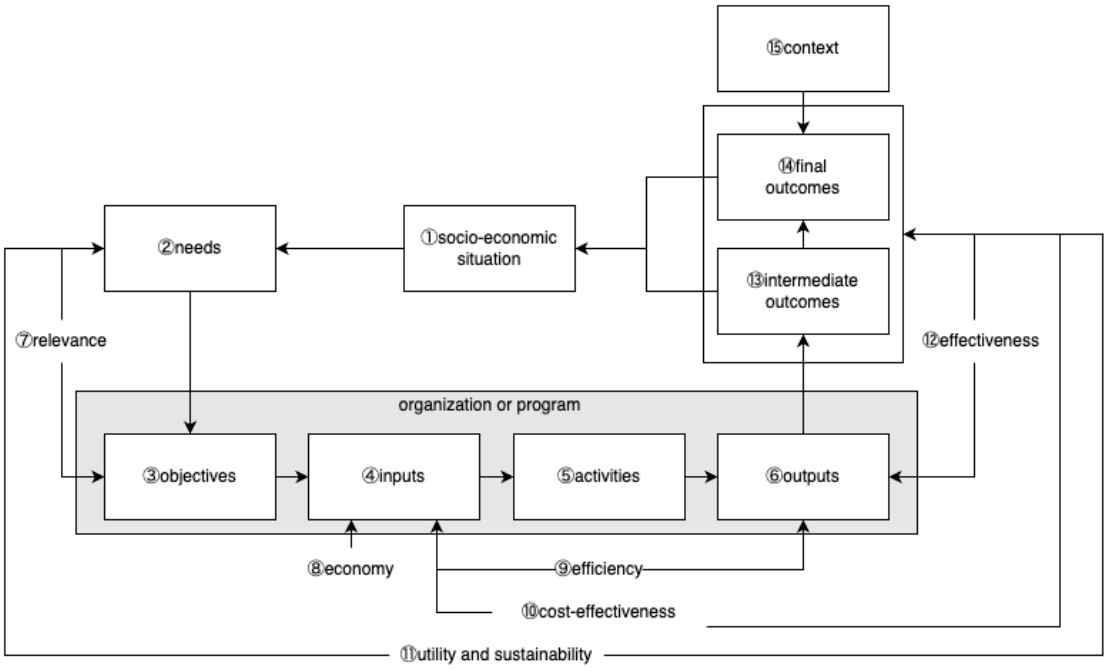



**Figure 2 Performance as a production process by Van Dooren et al., (2015)**

Within this process, *indicators* otherwise known as *performance information* (PI) are the measurements of the qualitative or quantitative subject(s) deemed important to the organization or program, and choosing PI comes from a process called *performance measurements* composed of five steps (Van Dooren et al., 2015). In this research, the keywords “indicators” and “performance information” will be used interchangeably. Performance management is the management of PI, and subsequently performance measurements, that is used in decision-making (Moynihan, 2008; Van Dooren et al., 2015).

### 2.1.2 Complexity of Performance

The complexity of performance is rooted in its subjective nature and becomes further complicated when discussing performance in a public sector context. First, performance measurements, and ultimately indicators, are selected by the organization or program based on their importance. Depending on this, methods for assessing performance can be selected but these methods may not reflect true intentions, nonetheless maintain objectivity (Boyne et al., 2003). Ultimately, subjectivity can trickle down starting from the organization itself to the indicators and methods, rendering results-based management unsuccessful in some cases (S. V. Thiel & Leeuw, 2002). Consequently, individuals and organizations assessing performance may reach out to relevant and diverse stakeholders to realize a "multiple constituency" method for an objective assessment, but factors like the selection of these stakeholders and what these stakeholders ultimately determine as good measurements of performance are also liable to subjectivity (Boyne et al., 2003, p. 15).

Second, assessing performance in a public sector context can be convoluted due to the nebulous objectives and scope of the sector. Performance in the private sector is more straightforward because the objective of the private sector allows for a one-dimensional view: profit maximization is the end goal, so performance can be measured through calculations such as productivity and efficiency of outputs, and effectiveness of outcomes using methods like cost-benefit analysis. Furthermore, performance in the private sector can be measured on various levels, but the highest level of analysis could be industry. On the other hand, the objective of the public sector is not profit maximization and the public sector itself is multi-dimensional, spanning the economy, health, society, and well-being (Mustea et al., 2021). Therefore, the units of inputs and outputs may not be represented financially or quantitatively as they usually are in a private sector context analysis, and the transformation of outputs for analysis is often contested (Goldstein & Spiegelhalter, 1996). Additionally, the "public sector" spans far and wide, and when considering scope, performance can be broken down into a micro, meso, or macro level (Van Dooren et al., 2015). Micro performance assesses individual organizations, while meso performance is a policy sector, a chain of events, or networks, and macro is the highest level that looks at local, regions, nations, and supranational organizations. Although micro performance assessments may be more easily defined, meso, and macro performance assessments become difficult to define and untangle from each other as they can overlap (Lovre et al., 2017; Van Dooren et al., 2015). The expansion of scope also applies when considering the sustainability context, as sustainability problems bridge micro and meso levels at the macro level (Shields et al., 2002).

Lastly, uncontrollable circumstances that impact the output (the result of activities) from its outcome (the impacts of the output) are prevalent in the public sector, making outcomes difficult to assess and identify. In fact, there are two types of outcomes: outcomes that are influenced, and outcomes independent of outside influences thereby the public sector has little control over (Van Dooren et al., 2015; Walker, 2002). This can also be seen in Figure 2 under "intermediate" and "final" outcomes. In theory, outcome measurements can lead to assessing the utility and sustainability of the organization or program against the needs born out of the socio-economic situations but considering the aforementioned two types of outcomes, it is difficult to accurately define intermediate and final outcomes especially when considering a highly complex problem and system like sustainability in society (Pham & Smith, 2013; Van Dooren et al., 2015).

Therefore, to better understand performance, it is imperative to lay out the context of performance in a theory-oriented way, and it will be done so two-fold through public value theory to establish context, and the agency stewardship theory to establish the perspective of performance assessors.

### 2.1.3 Context to Performance: Public Value

Although the process of performance is known, the exact definition of performance can be elusive. Thus, Van Dooren et al. propose examining public value as the context of the situation that is being evaluated thereby giving shape to performance (2015, p. 28). There is no concise definition of "public values" in public value research, and prominent researchers base their definitions on different ideals. For example, Mark Moore's economical take on public values pins managers of the public sector to fulfill citizens' desires efficiently and economically (1995), while Barry Bozeman's public value is based on the provision of normative consensus that leads to the creation and common understanding of public values (2007). Bozeman has also created a "public value failure criteria" framework that identifies ten types of public value failure in markets that are deemed efficient but do not provide public value (2002).

Understanding public values as the context of performance is vital, as public values can help contextualize the performance and its characteristics while adversely, performance can be seen as the realization of public values. Ideally, what is being evaluated in the performance process aligns with the societal needs that had created the policies or programs in the first place. In that way, public values can aid in the assessment of performance (Van Dooren et al., 2015). If one considers the influence of public values on the variable "context" on outcomes as seen in Figure 2, then it is imperative to also take into consideration the defining public values of the context to assess performance. Thus, it is difficult to generalize what constitutes inputs, outputs, and outcomes in the public

sector context because of context-dependent definitions of public values. Therefore, to operationalize inputs, outputs, and outcomes appropriately and accurately, it is vital to look at the context and the underlying public values.

### 2.1.4 Context to Performance: Influences of Agency and Stewardship Theory

As previously mentioned, one of the complexities of performance is its subjectiveness, and in assessing for subjectivity, the actors involved in the performance process and their intentions should be examined, especially as LRGs may be controlled or steered by their principal government in the performance process. A key theory is the agency theory in which agents are opportunistic in the face of their principal with two kinds of consequences as a result of information asymmetry between the two actors with the agent having the advantage: adverse selection and moral hazard. Agents can purposefully mislead principals through adverse selection ex-ante a contract between the two actors. For example, to gain the principal's approval, agents may modify performance measurements to embellish the efficiency and effectiveness of the agents' action, and this process is labeled as "measure fixation" (S. V. Thiel & Leeuw, 2002). Additionally, because of the opportunism of the agents, moral hazard can arise as an ex-post consequence of "imperfect monitoring" as a reflection of the agent's deterrence from the principal's goals (Verhoest, 2017). It must be noted that in the context of the SDGs, a similar concept to agency theory – multi-level governance (MLG) – is a core part of it especially as the implementation of the SDGs requires both horizontal and vertical cooperation and integration within governments and society that is central to MLG. However, for the scope of this research, agency theory will be taken into consideration.

To adjust for these risks, the principal may control the agents through four control mechanisms: input control, control by output or results, results-related financing, and introducing competition (Verhoest, 2017, pp. 11–12). More generally speaking, these four control mechanisms can be categorized as 1. efforts by the principal to monitor the agent's performance, 2. bond arrangements, or 3. establishing incentives and risk-turnovers that ultimately influence the performance of the agent through goal alignment, and motivation, and behavioral control. For example, the control of input via the creation of rules and procedures may greatly reduce the decision-making power of the agent and can be used to combat moral hazard. Other methods such as controlling results through setting standards and targets, linking results of agents to budgets through performance budgeting, and benchmarking performance to competition all allow for the principal to counter the agent's opportunistic behavior. However, observing agency theory alone may not fully account for agency behavior, as pointed out by stewardship theory scholars.

As opposed to the hierarchical control of agency theory, stewardship theory reflects on goal congruence and a trust-based relationship between the principal and agent (Bjurstrøm, 2020). Although they are not dichotomous, the difference has implications for how and what control mechanisms of performance may work depending on if a party is an agent or a steward; if the relationship resembles agency theory, then the aforementioned strict controls should be utilized but if the relationship resembles stewardship, then having strict controls would be fruitless (Bjurstrøm, 2020). The controls of a principal-steward relationship are similar to the controls of a principal-agent relationship, and it can be conducted via bond arrangement or contracting, or process management of performance. In terms of contracting, agency theory suggests a "complete" contracting method with as many contingencies, objectives, and targets being covered and specified that ultimately results in strong control while steward theory suggests a "relational" contracting method where there is less control. In terms of process management, agency theory shows a top-down approach to managing performance where the principal sets goals and performance targets and utilizes performance management for control rather than learning or for accountability as previously mentioned. On the other hand, steward theory shows a bottom-up approach to performance management where the stewards have greater autonomy, goals, and performance targets are created jointly by the principal and the steward and utilizes performance management for learning rather than control or accountability (Bjurstrøm, 2020).

A summary of the agency stewardship theory and its implications on performance management in terms of control mechanisms is shown in Table 1.

**Table 1 Summary of Implication of Agency and Stewardship Theory on Performance Management by Bjurstrøm (2020) and Verhoest (2017)**

| Type of Relationship | Main characteristic | Control mechanisms | | |
|---|---|---|---|---|
| | | 1. Monitor performance | 2. Bond/contract | 3. Establish incentives |
| Agency | Based on agency opportunism as a result of info asymmetry | Top-down, principal creates rules, procedures, standards and targets | "Complete" contract with hard goals | Performance-based budgeting, market financing and competition |
| Steward | Based on goal congruence and trust | Bottom-up, collaborate on goals | "Relational" contract with flexibility | No incentives as the autonomous relationship is based on trust |

## 2.2 Conceptualizing Performance Measurements and Indicators

The definition of performance can be elusive and largely relies on the organization or program assessing it. What does this imply for indicators? What are indicators, how do they come about, and what are the potential pitfalls of using indicators? These are the

questions that will be answered in this section which is guided by the steps of the performance measurement process.

### 2.2.1 Indicators within the Performance Measurements Process

According to Van Dooren et al. (2015), performance measurements are also a process comprising five steps: 1. Prioritizing what to measure, 2. Selecting indicators, 3. Collecting data, 4. Analyzing, 5. Reporting (2015, p. 63). These five steps are further supported by "quality assurance" (2015, p. 63). It is important to consider this entire process in order to truly understand the significance of the indicators themselves.

First, the priorities for measurement must be laid down due to the limitations in resources and capacity of the public sector. Van Dooren et al. (2015) state three key questions to define priority: the context of the measurement, priorities for measurement, and reasons for priorities (2015, p. 66). Context can be understood using tools such as management models, trees of objectives, stakeholder analysis, and program logic. This research will also consider the aforementioned contexts of performance (public value, agency stewardship theory) as part of contexts of measurement. Priorities for what to measure and the reasoning for those priorities will largely be influenced by how PI is expected to be used, which is mainly in three ways: to learn, to steer and control, and to give account (Van Dooren et al., 2015, p. 120).

### 2.2.2 Typology of Indicators

The next step in the performance measurement process is the indicator selection. But first, the definition of indicators and what they are will be explored. The term "indicator" means "one that points out" (Merriam-Webster, 2023). The OECD defines indicators as "contextualized information for specific purposes" (2008, p. 8). For example, a needle on a speedometer in a car displays the indicator that is kilometer per hour (Gudmundsson et al., 2016). They play a pivotal role in representing measures that can then be used for analysis. The types of indicators are plentiful (Brugmann, 1997) and for the scope of this research, four main types of indicators will be discussed: dimension, complexity, position, and stage indicators (Gudmundsson et al., 2016). A summary of the four types of indicators is seen in Table 2.

**Table 2 Typology of indicators by (Gudmundsson et al., 2016)**

| Type | Main characteristic | Sub-types |
|---|---|---|
| Dimension | Time and space (place) of measurement | • Single point<br>• Time series<br>• Cross-section<br>• Cross-section and time series |

| | | |
|---|---|---|
| Complexity | Content of measurement | • Descriptive<br>• Ratio<br>• Normative<br>• Index |
| Position | Ex-ante or ex-post analysis | • Leading<br>• Lagging |
| Stage | Phase within the performance process that the measurement takes place | • Input<br>• Output<br>• Outcome |

*Dimension* indicators reflect the time and space (place) of the measurement and can be broken down into four types: 1. single point, 2. time series, 3. cross-section, 4. cross-section, and time series. A single-point indicator reflects one variable, at one place, at one time, while a time series indicator reflects a comparison over time. A cross-section indicator reflects a comparison of entities such as cities and groups, while the cross-section and time series are a mix of the second and third types.

*Complexity* indicators reflect what the indicator is communicating in terms of four types: 1. descriptive 2. ratio 3. normative 4. index/composite. Descriptive indicators are the "illustration of a condition using a particular variable" while ratio indicators are used to assess "relative improvement" by dividing at least two variables to calculate a ratio. Normative indicators are assessed against a reference point whether that point is determined legislatively with the existence of regulations and standards, scientifically, or voluntarily (Gudmundsson et al., 2016; Kaufmann & Kraay, 2008). Index indicators, otherwise known as composite indicators (CI), are an aggregate of sub-indicators as opposed to the other indicators stated thus far which are considered "single indicators" that reflect one measurement of one objective (OECD, 2008). CI is used when measuring complex and multi-dimensional topics, such as the contested gross domestic product, and allows for comparisons (GDP; Laurent, 2018). Therefore, composite indicators are often used when examining sustainability but its usage and mathematical complexity call for careful scrutiny when using CI as it could potentially be used to create narratives for policymakers rather than objective measurements of performance (Kuc-Czarnecka et al., 2020; OECD, 2008). As composite indicators are a summary of single indicators, one small adjustment can skew the overall results drastically. This simplification of information can sometimes be calculated based on unclear math, coined as the term "mathiness" by Nobel prize winner economist Paul Romer, to purposely mislead results for political agendas which Romer himself was a victim of as he had to step down as Chief Economist of the World Bank after a miscommunication over CIs (Kuc-Czarnecka et al., 2020; Romer, 2015). Furthermore, creating CI relies on the value judgment of the organization of every single indicator, thereby giving the freedom to the organization to pick and choose single indicators that may create a better CI result (Bird et al., 2005). Therefore, there are various guidelines and recommendations to prevent these pitfalls.

One of these dominant guidelines in practices is a handbook by the OECD and Applied Statistics and Econometrics Unit of the Joint Research Centre (JRC) of the European Commission published over a decade ago that spells out creating CIs in ten steps: establish a theoretical framework, select data, input missing data, multivariate analysis, normalization, weight and aggregation, uncertainty and sensitivity analysis, back to the data, links to other indicators, and visualization of the results. Ensuring the quality of each step is critical, and breaks down quality into seven dimensions: relevance, accuracy, credibility, timeliness, accessibility, interpretability, and coherence (OECD, 2008). In academic literature, it is found that researchers will utilize the OECD guidebook for creating CI in the analysis (e.g. Reggi et al., 2014).

*Position* indicators reflect an ex-ante or ex-post time description of the situation by the indicators; if forecasts can be made by the indicator, they are considered "leading" indicators while "lagging" indicators confirm past events. For example, test scores may be considered a lagging indicator that suggests future growth (May & Sanders, 2013). It is said that ascertaining an indicator between leading and lagging may be tricky, as deciphering it requires the context of the situation as well as an understanding of the specific cause-and-effect cycle that is difficult to establish in and of itself.

Finally, *stage* indicators reflect the stage in the process of the context. In this case, the context can refer to a policy or a program, as seen in Figure 2, and so there are three main stage indicators: input, output, and outcome indicators. Input indicators, otherwise known as process indicators, are seen as raw materials e.g. finances, human resources, while output indicators are the result of activities, and outcome indicators are the impact. For example, to assess a country's productivity level, the input indicator can be set as "human capital" and the output as "gross domestic product per worker".

In addition to the types of indicators, there is a multitude of research on what makes an indicator 'good'. The SMART framework is often used in both public and private sector contexts; they should be specific, measurable, achievable, relevant, and timely (Gudmundsson et al., 2016). This framework was, in fact, used to guide the creation of the global SDG indicators (SDSN Secretariat, 2014). However, there are other factors that are important to consider especially in the frame of the public sector and wicked issues. For example, Van Dooren et al. (2015) describe six characteristics of good indicators: sensitive to change, precisely defined, understandable for users, documented, relevant and actionable, feasible data collection, and compliant with coordinated data processes and definitions (2015, pp. 69–70). Bouckaert (1993) mentions 20 criteria for valid performance measurements, including countable, uniform over time, and accuracy (Bouckaert, 1993, p. 32). However, Bouckaert acknowledges that valid performance

measurements can be easy to assess for tangible outputs and outcomes such as garbage collection, while it becomes more difficult for "human"-related activities such as education, and "idea"-related activities such as think tanks and policy staffs. Bird et al. (2015) list 14 definitions for indicators that aid in defining PI for achieving PM's objectives, including being directly relevant to PM's primary objective, precise but practicable, consistent, uncertain, and others (2005, p. 6). A summary of the qualities of "good" indicators can be seen in Appendix 1.

### 2.2.3 Use of Indicators and Performance Measurements by Stakeholders

There is various research into indicator usage. For example, performance information are said to be for *guidance* including preparation and decision, and *control* including execution and evaluation (Bouckaert, 1993, p. 34). Additionally, indicators can be used as *targets* that specify standards and thresholds, *rankings* that allow for comparisons and *intelligence* to gather background information (Hood, 2007). For this research, indicator usage as stipulated by Van Dooren will be the main definition taken into consideration but parallels to Hood's definitions will be explored as well.

Van Dooren states three main usages: *to learn* or for research purposes, *to steer and control* or for managerial purposes, and *to give account* or for democratic purposes (Bird et al., 2005; Van Dooren et al., 2015). If PI is to be used for learning, which is comparable to Hood's intelligence usage, then factors that relate to learning such as facilitating dialogue, developing leadership, and changing culture are needed (Moynihan, 2008; Van Dooren et al., 2015). However, Hood warns that this kind of usage may risk a "lack of transparency and clear incentives", as indicators may be subject to interpretation due to the lack of targets that would further demote public managers to not pursue indicators (Hood, 2007). If PI is to be used for steering and controlling that is based on "keeping track", then using management scorecards like a balanced scorecard (BSC) or policy monitors should be used. If PI is to be used for accountability, then communicating performance as a way of understanding and justifying past performance is needed. Hood's usage of targets and rankings can be applied to Van Dooren's usages of steering and controlling and accountability, and in that case, Hood and other researchers warn of the pitfalls of rachet and threshold effects, and the focus on rankings that risks distortion of results and ultimately misrepresents reality (Hood, 2007; Van de Walle, 2008).

The implications of these three uses along with their methods are tied to the "soft" and "hard" usage of indicators (Van Dooren et al., 2015). As seen in Figure 3, plotting them on these axes may make it easier to understand why the aforementioned methods for each use are appropriate. Indicators can be coupled to judgment, whether tightly or loosely, with the implication of assessing performance to results or allowing room for dialogue

respectively. Consequently, this coupling affects the impact it has on the organization, with tightly coupled indicators tending to have a higher impact and vice versa (Van Dooren et al., 2015). Despite its 'strictness', hard use that is tightly coupled to judgment may give way to manipulation, as evaluations may occur through rankings and comparisons to targets that may be cherry-picked. Furthermore, there is a higher risk of errors, ranging from simple calculation errors to sampling errors and categorization errors(Hood, 2007). On the other hand, soft use may risk a kind of unpredictability, specifically because of the ability to have multiple interpretations of the results (Hood, 2007).

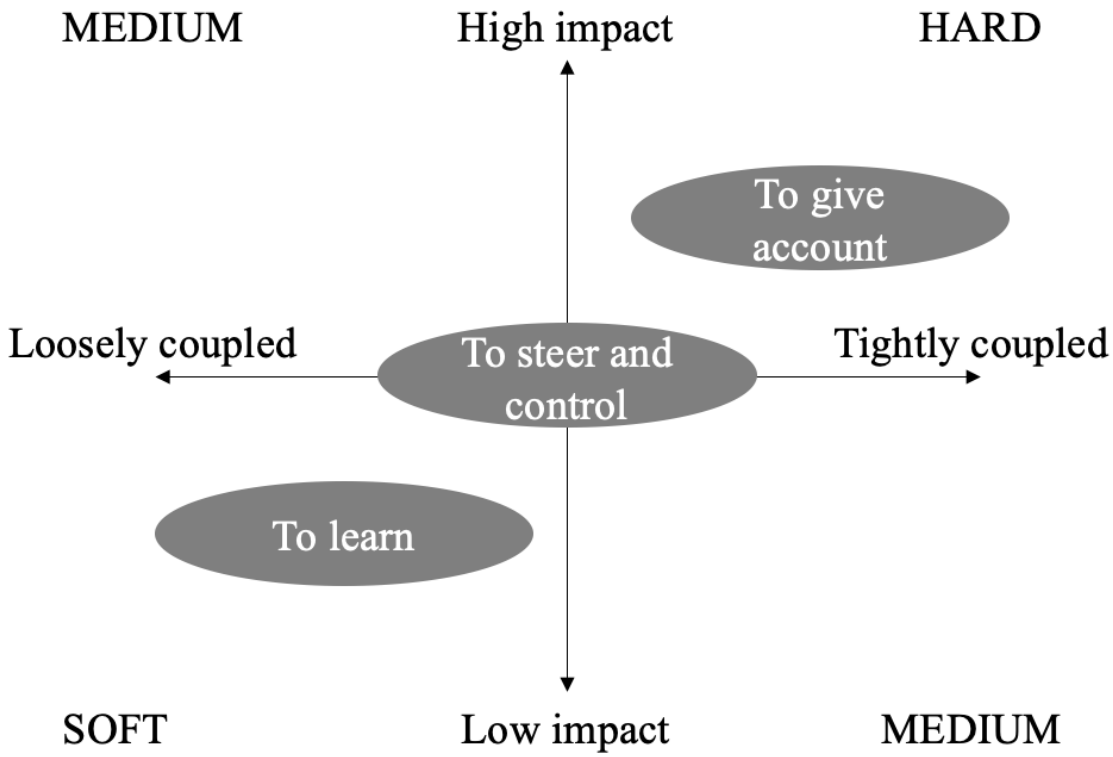



**Figure 3 "Assessing the nature of use of performance information" (Van Dooren et al., 2015, p. 124)**

However, regardless of the specific way PI is used, it can be said that incorporating it in decision-making is the sole way of defining whether performance-oriented reforms are successful (Kroll, 2015). In other words, performance management can be deemed a success only if PI is genuinely used.

### 2.2.4 Potential Pitfalls of Indicators: Subjectivity and Impartiality

It has already been made clear that the concept of performance can be subjective to the organization or program, and indicators are no exception to this. There are several pitfalls

of indicators that have been the foci of research, starting with its subjectivity. The concept of indicators with its basis in NPM may overlook difficult-to-quantify "socially important dimensions" during the quantification of progress when used in the public sector (Swianiewicz, 2020). This quantification may be at the mercy of policymakers and their individual agendas (Allen et al., 2019; Bird et al., 2005; Spillane, 2012). For example, public servants delivering public services may use PI to find gaps for improvement of the specific service, and managers may use PI to assess the quality of and create a basis for the organization. Board members may use PI to hold managers accountable. Citizens may use PI to assess personal satisfaction with public organizations and services, and this ultimately can influence citizen trust in public organizations as shown in empirical research where the publication of PI led to higher trust (Yang & Holzer, 2006). Therefore, indicators may not represent the complete reality but a carefully selected fragmentation that reinforces political agendas. All of this may occur because of various assumptions such as the origins of indicators, or data, being neutral, or that indicators are created independently of any processes that involve stakeholders' behaviors (Bird et al., 2005; Hasselblad, 2021). This highlights the importance of processes and context to indicators.

Some research points to the importance of agency-level actors instead of central officials and legislators in PI due to the scope of their daily work. Agency-level actors are "more specialized and more homogeneous" thus they are more likely to understand which of the PI are relevant and be more motivated to learn (Moynihan, 2008, p. 197). The implications of this influence how performance management should be organized thereby redefining the role of the aforementioned actors. If PI were to be concentrated at an agency-level, then it can be the role of the central bodies to support agency-driven change rather than enforcing compliance with PI requirements to those agencies and vacuuming massive amounts of data from them (Moynihan, 2008). However, it is also important to protect the impartiality of "compilation, presentation and statistical interpretation of performance measures" as the possibility of regarding results being skewed by the public is possible and this could result in a decrease of trust and confidence (Bird et al., 2005, p. 11). To counter these possibilities of impartialities, the involvement of third parties to oversee the processes to create a more "democratic" process in oversight, or setting high standards to ultimately create objective knowledge of policies arise (J. Heinrich, 2007). In conclusion, it is critical that the role and responsibilities of the municipalities or cities are analyzed, including which part of the performance process lies in their scope, how the data is collected and analyzed, who compiles the results, and who and what is audited throughout the whole process.

### 2.2.5 Data Collection

The main characteristics of data collection are availability, aggregation, and sources(Van Dooren et al., 2015). These characteristics will be explained below.

*Data (un)availability*

Data availability may influence what is chosen to be an indicator, as this point is taken into consideration when initially coming up with indicators. At the same time, the unavailability of data also is a challenge undermining many public organizations as indicators should not be prioritized when data is not available, but this does not reflect the unimportance of that indicator.

*Data source*

The variety of different sources of data can influence the quality of data, as well as cost (Van Dooren et al., 2015). In terms of quality, collecting data using expert assessments versus, for example, firm or household surveys may result in different results; although expert assessments may reduce overall cost as there are fewer experts than firms or households on a given topic, their expertness may prove a hurdle as biases may interfere in results (Kaufmann & Kraay, 2008). Using administrative data, such as data collected routinely by public institutions, also reduces cost but oftentimes, the collected data do not directly measure the data needed by the indicators (Merry, 2019). Instead, what the collected data can do is act as a proxy, but this may feed into the bigger problem of a "proxy problem" in which the provided data do not directly answer the target thereby potentially producing misleading indicators (Merry, 2019).

*Aggregation of data and data sources*

When discussing aggregation of data, sampling versus complete enumeration as well as the aggregation of data from different sources are questions that must be asked. Sampling is choosing a set *n* to represent the entire population while a complete enumeration is a thorough data collection (Bird et al., 2005). If one were to conduct sampling, then it is imperative to have an audited and specified sampling procedure. In some cases, sampling may result in better quality data while reducing costs and predictability of the data collection (Bird et al., 2005). However, the notion of sampling on a greater scale is especially debated in the context of sustainability as it may marginalize certain groups of people thus not depicting a whole and thorough picture, and there is a general trend to "disaggregate" data in terms of "income, sex, age, race, ethnicity, migratory status, disability, and geographic location, or other characteristics" (Inter-Agency and Expert Group on & Sustainable Development Goal Indicators, 2019). In addition to the sampling

of data, aggregation of sources is also a challenge similar to the aforementioned explanations. Therefore, a mix of sources is considered a good method of triangulating results to ultimately provide a more accurate picture of the result.

### 2.2.6 Data Analysis

Third, the *analysis* of these data points is conducted to transform data into usable information (Van Dooren et al., 2015). Some researchers propose specific ways how data analysis can be conducted, such as with norm and target setting, breakouts, and causal analysis (Van Dooren et al., 2015, p. 73). This definition of data analysis can be thought of as the *methodology* that acts as a frame and compass for the overall study which is performance management. The selection of a *method*, or tools of data collection and analysis, can be viewed as reflecting the bias, views, and assumptions of a researcher on reality, which can be connected to fundamental concepts of philosophy such as objectivism that aim to objectively predict and explain people, phenomena, and systems and others that go beyond the scope of this thesis (Moon et al., 2019). Therefore, it is said that research should describe the chosen methodology including any philosophical contexts that answer how they collect, analyze, and interpret the data as it forms the basis for the method (Wolcott, 2008).

By understanding the methodology, it becomes easier to determine if the chosen methods are appropriate. For example, Van Dooren et al. state the method of breakout analysis may be based on hypotheses that reflect the researcher's subconscious biases that otherwise may be reflected in the methodology, thus if the methodology is clear, other kinds of methods such as interviews and focus groups could supplement the breakout analysis for a fair analysis (2015, p. 76).

### 2.2.7 Reports

Fourth, the findings of these analyses as well as the process that leads up to that analysis including information of context are put together in a *report*. When compiling such a report, it is important to identify the main stakeholders who will be consuming the report as it can influence the communication medium and how the information should be presented. Mediums can include websites, social media, letters, financial documents, and other types but for this research, reports will be examined in Section 2.3.

### 2.2.8 Quality assurance

Lastly, each of the above five steps should be supported by *quality assurance* of the performance information and the measurement system (Van Dooren et al., 2015). This is

to ensure that non-use, misuse, or over-use of performance information does not happen, as these can have negative effects on policymaking that may ultimately corrode the trust of stakeholders and citizens (Lyytimäki et al., 2013). Quality assurance can be applied to both the statistical data as well as the entire process. For statistical data, multiple layers of control systems should be created for quality control. To be more specific, Van Dooren et al. recommend an internal control system and internal audit within an organization, and an external audit independent from the organization to audit the data (Van Dooren et al., 2015, p. 81). The quality assurance of the process can be confirmed through the adherence of the process to the aforementioned characteristics of defining priorities, selecting indicators, collecting data, conducting analysis, and creating reports.

Through this process, the validity, legitimacy, and functionality of the measurement systems can be assessed to determine "fit for use" (Bouckaert, 1993, p. 31). Validity, as the main priority as well as the "traditional focus for good measure", is the logical soundness and robustness of "a mechanism, a theory, a system… and a classification", and is composed of the 20 characteristics of valid measurements as explained in Section 2.2.2 (Bouckaert, 1993). Included in validity is the concept of "reliability", or repeatability. Functionality, or if the measurement is fit for use, can be broken down into non-functionality and dysfunctionality, in which the former refers to the ignoration of information while the latter refers to "negative effects due to measurement" (Van Dooren et al., 2015, p. 81). Finally, legitimacy refers to the ownership or general support of organization members of the measurement system.

To echo the aforementioned points, other scholars have brought up vital values necessary throughout the performance management process: *integrity*, *confidentiality*, and *ethics* (Bird et al., 2005, pp. 23–24). Performance management can be utilized to hold public organizations accountable, thus it must be ensured that the process of performance management itself has integrity; the same stakeholders who define performance information (PI) of accountability should not be assessing it for themselves. Moreover, the statistical integrity of calculations must be considered, as "major corruption of measured indicators" could occur in the worst-case (Bird et al., 2005, p. 23). Lastly, ethical concerns of performance management arise as PM may use data about citizens, thus problems of consent and data protection of citizens must be discussed as well as defining legislation of personal data access for PM. These points are made aware by the UN, as they created a nine-point guidance note on the use of big data and thereby the guarantee of data privacy, protection, and ethics including "lawful, legitimate and fair use", "purpose specification, use limitation and purpose compatibility" and "risk mitigation and risks, harms, and benefits assessment" along others (United Nations, 2017a).

The performance measurement process and quality assurance represent an ideal; it may not be that all components are realized in reality nor occur in this particular sequence. However, Van Dooren et al. (2015) state that having this ideal type can be used to identify deviations from this "pure measurement model" (2015, p. 63), and therefore this process and the quality assurance component will be considered when assessing indicators and reports.

## 2.3 Conceptualizing Reports in the Public Sector

The existence of performance management alone does not improve the effectiveness of an organization; one research found the leading cause of effectiveness for change within an organization was the *communication* of the said performance management system (Cunningham & Harris, 2005). Reports are one way of fulfilling this communication role, and thus in this section, reports in the public sector will be examined. First, the general functions and roles of reporting in the public sector will be introduced. Then, the key characteristics of non-financial reports and information as well as the role of those reports will be examined. Although financial reports continue to be pivotal in terms of understanding public management, this research will focus on non-financial reports as sustainability reports are separate from financial reports. After, reporting guidelines and standards currently being used will be introduced as well as its academic critique. Finally, the strategic disclosure of information in these reports will be examined, whether it be through graphical representations or other key features of a report.

### 2.3.1 Function and Aim of Reports

Public reporting, or the publication of information via reports, is defined as "a government-led process of collecting, analyzing and communicating quantitative and qualitative data" (Bexell & Jönsson, 2019). It can be thought of as a post hoc "report card" that reflects on a public organization's activities (Lee, 2002). Reporting can either be mandatory or voluntary. Mandatory reporting is bound by laws, such as the Corporate Sustainability Reporting Directive of the EU that specifies the content and these reports are under the subjugation of audits. Voluntary reporting can be either entire reports that are not required by law or are voluntarily-added parts to mandatory reporting that add characteristics like narratives (Clarke et al., 2009; European Commission, n.a.). The case studied for this research are classified as "voluntary" (UNDESA, 2020).

According to Lee, there are two methods of reporting: indirect via media or direct via the public institution or organization (2002, p. 37). The media, with its widespread influence on the public, acts as a watchdog and monitors governments and their officials (Francke, 1995). However, Lee notes that media coverage of public administration is decreasing

thus the importance of direct reporting, whether it would be in the form of annual reports, Internet websites, electronic chat rooms, or public access channels, is evermore increasing (2002, p. 39). Annual reports, such as financial statements, are particularly gaining attention as the main medium of conveying government accountability, especially on the local government level where the needs to be transparent yet personal are more prominent than on other levels (Roundy et al., 2022; Ryan et al., 2002; Steccolini, 2004).

Direct reports can be classified as financial, non-financial, or a mix of both otherwise called "integrated reporting". There is a large amount of research on financial reporting due to its long-standing practices, but comparatively less on non-financial reporting, especially in specific contexts like improving public sector performance (Budding et al., 2022; Farneti & Guthrie, 2009; Ogata et al., 2018). Nevertheless, non-financial information and reporting are greatly used and prioritized by both politicians and managers as financial information alone cannot capture the full extent of public activities and their outcomes (Farneti & Guthrie, 2009; Liguori et al., 2012; Montesinos & Brusca, 2019; Van Dooren et al., 2015). Additionally, the growing demand for transparency and accountability in public organizations, and a loss of trust by citizens contribute to the need to showcase performance beyond non-financial means (Montesinos & Brusca, 2019). Furthermore, non-financial reports have expanded beyond their traditional scope of measuring policy outcomes, and into areas such as environmental protection and social responsibility information as public value has changed (ECIIA & EUROSAI, 2021). Therefore, it can be said that non-financial reports are expected to be catered to certain stakeholders and acknowledge specific demands, provide context for performance, and communicate in a transparent and accountable way to win back trust (Montesinos & Brusca, 2019). Nonetheless, a lack of initial standards and regulations for non-financial reporting has led to the birth of various styles of reports, mainly sustainability reports (SR), integrated reports (IR), and others that will be described below (Biondi & Bracci, 2018).

The origins of the scope changes in non-financial public reports can be traced to the private sector, where industries were expected to work beyond its profit maximization objects and to areas such as social responsibility and with that, the need to report on these non-financial activities (Montesinos & Brusca, 2019). The private sector had birthed social responsibility reports, which reflect the actions of corporate social responsibility (CSR) or the obligation of businessmen to conduct activities in line with societal values and objectives (Latapí Agudelo et al., 2019). Additionally, sustainability reports (SR) are a way to incorporate societal and environmental information within their pre-existing financial reports (Montesinos & Brusca, 2019). Due to the aforementioned initial lack of standards or regulations for these reports, there is no universal definition for these kinds

of reports (Farneti & Guthrie, 2009). Despite its ambiguities, it is noted that CSR and SR reports of the private sector are now made mandatory in the European Union under the Corporate Sustainability Reporting Directive of 2023 (European Commission, n.a.). The recency of this Directive sheds light on the ever-increasing pressure felt across all corners of society to produce reports regarding sustainability and performance.

*Aim of reports*

Public reports have various aims such as proving *organizational legitimacy*, holding the organization *accountable* through reports that provide *transparency*, and serving as a tool of democracy. Unlike private organizations or companies that are at the mercy of public institutions and public standards, there are fewer measures available to check and control public institutions. However, it is still vital for public institutions to have legitimacy in order to fulfill organizational objectives that may otherwise not be executed by other institutions and organizations. *Legitimacy theory* implies the existence of outside forces that shapes the ways public institutions show legitimacy, such as government mandates, media, and legislations, and the existence of a "social contract" in which the public institution is expected to fulfill to society (Guthrie et al., 2004). In order to showcase compliance with these outside forces, public institutions craft reports using standards and systems to create a narrative, sometimes using impression management that may generate a more favorable perspective of these institutions (Samkin & Schneider, 2010). Another definition commonly used in the specific context of sustainability and environmental reporting of *legitimacy theory* is attaining legitimacy by aligning "social values and organizational behaviors" (Lodhia et al., 2012). The main concern for organizations, in this case, is plugging in a *legitimacy gap,* or the misalignment between social values and organizational behaviors, and it was found in previous research of private sector firms' sustainability reporting that this fear of the legitimacy gap led them to frame itself in a positive way through the use of impression management. Therefore, acting upon organizational legitimacy has been scrutinized as a way of controlling reporting (C. Edgar et al., 2018). In this way, reports can create a façade to deflect differing stakeholder demands (Bexell & Jönsson, 2019), while reaffirming their compliance with standards.

Accountability is another key concept related to legitimacy that can also be seen as an outcome of reports that provide transparency. Accountability is composed of two main aspects: doing what the organization said it would, and being responsible for it (Stefanescua et al., 2016). Reports as a way of showing the results of its actions is a key way of showing these aspects, but the existence of a report does not automatically translate into increased public accountability (Cunningham & Harris, 2005). Therefore, using performance measures in addition to providing these reports could be crucial in

some contexts. On the other hand, reports can be interpreted as a reflection of transparency efforts, or the "availability of information about an organization that allows external actors to monitor the internal activity and performance of that organization", thus leading to public accountability (Stefanescua et al., 2016). Previously, public transparency was limited to mainly a financial scope such as taxes, but now it has extended to areas such as sustainability and through media such as reports and websites that are supported by international guidelines that recognize a need for transparent information that further supports governments and public administrations to publish a 'good' transparent report (Stefanescua et al., 2016).

In addition to legitimacy, accountability, and transparency, reports also serve as a tool of democracy. Reports inform the public about a government's activities and operations, thereby fulfilling a fundamental function as the backbone of democracy (Lee, 2002). By informing the public, citizens can hold public opinions that fuel discussions on important matters to society. The importance of informing the public stems from the obligation of public institutions to be accountable as part of a functioning democracy (Bexell & Jönsson, 2019). Reports reveal not only the performance of public institutions but also the fact that they are acting within various standards and practices of different institutions (Lee, 2002, p. 36). However, it may be these very standards and practices that can be used as a gateway for these public institutions to strategically reveal certain information to influence readers, as will be explained in the following section.

### 2.3.2 The Role and Impact of Text and Visual Representations in Public Reports

Reports are composed of textual and visual representations, or "graphical display of abstract information", and there is abundant research on the influence and positive role of visual representations that influence behavior and preferences, and encourage deeper understanding and information usage (Ballard, 2020; Choi & Gil-Garcia, 2022; Eberhardt & Silveira, 2018; Tufte, 2001). One study researched the influence of specific visualizations i.e. compositional formats including bar charts, pictographs, non-numeric and numeric tables, as well as the content including ordering, visual explanatory cues, amount of content, instructive aids, and representation of uncertainty against the three decision-making domains of comprehension, choice and preference (Hildon et al., 2012). Considering the user's perception abilities, tables and pictographs aided decision-making better than bar charts as bar charts were harder to understand. However, pictographs can overestimate harm, and bar charts were preferred when multiple risks were shown. Additionally, ranking and ordering aided in faster comprehension, visual explanatory cues were welcomed as they clarified the data, and the saying 'less is more' applied to

the amount of content as more content could lead to information overload (Hildon et al., 2012).

In addition, visual representations such as graphs are a good way to objectively catch the reader's attention, and identify trends in an accessible form but can also be carefully chosen based on the report creator's intentions (Jones et al., 2018). Creating graphs of good news was found to lead to a more favorable impression of the organization, and an analysis of current annual reports exhibited "selectivity, graphical measurement distortion, and use of presentational enhancements" (Cüre et al., 2020; Zhang, 2020). Utilizing visualizations in this way can be referred to as *strategic disclosure* as the report creator is intentionally choosing the pieces of information to visually represent, and the process of doing so is called *impression management* (Cüre et al., 2020; Melis & Aresu, 2022). This practice of impression management is especially common in private-sector companies, where the use of impression management ultimately made sustainability reports "about fostering positive public relations than providing a meaningful accounting of the social and environmental impacts of the firm" (Cho et al., 2012a; Jones et al., 2018).

There are two main types of impression management: enhancement for highlighting the positives, and obfuscation for hiding and misleading bad news (Cho et al., 2012b; Cüre et al., 2020). Visual examples of obfuscation and enhancement representations can be seen in Figure 4.

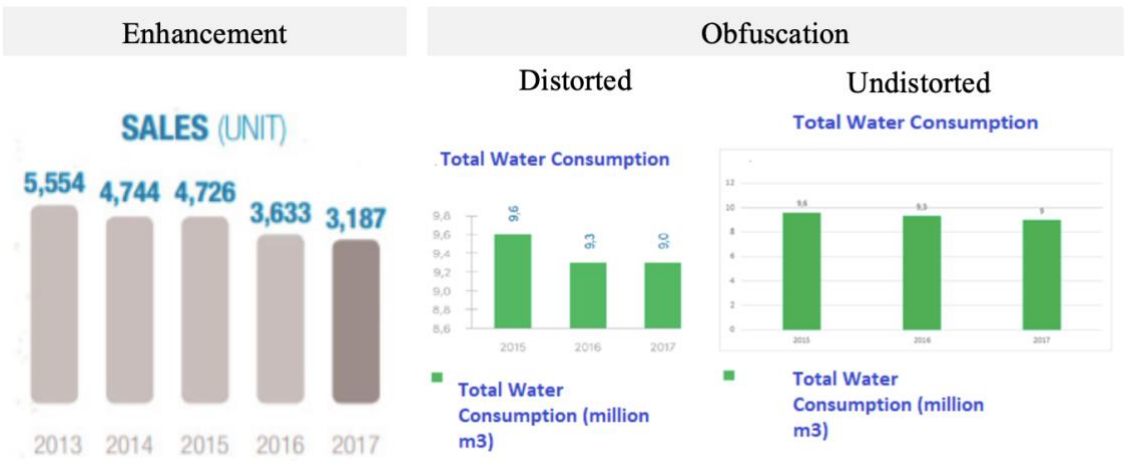



**Figure 4 Example of enhancement and obfuscation in visualization (Cüre et al., 2020)**

For enhancement, the bars of the graph are not visually accurate; the visual representation of the 3,187 bar from 2017 makes it appear not as significant of a decrease from its 2013 figure of 5,554, when in actuality, the visual representation for 2017 should be around

1.5 times smaller. For obfuscation, the distorted graph starts from the value 8.6 which allows for the bar graphs to show a greater difference than if the bar graph were to start from the value of 0, as seen in the undistorted graph to the right. To control impression management, Zhang recommends regulators lay out guides or limits on graphical representations to suppress their effects on the reader's behavior (2020, p. 247). Through these kinds of methods, the organizations in question in past research were shown to manipulate how readers of these reports would instinctively interpret these results.

On top of these known tactics, one characteristic specific to public sector reports is their textual complexity thus risking the inability to relay information appropriately; at the expense of public institutions trying to be as clear as possible, they may confuse their readers including citizens and oversight bodies (Roundy et al., 2022). This may be influenced by the public organization's efforts in educating the general public about sustainability (Manes-Rossi et al., 2020). To counter this risk of incomprehension, Roundy et al., recommend uploading reports in a machine-readable form rather than a scanned image (2022). Furthermore, Bird et al. recommend any performance information intended for publication should have an "intuitive appeal" (2005, p. 18).

### 2.3.3 Report Guidelines

There are international guidelines that aid public and private organizations in formulating reports. The main international guidelines are offered by the International Council for Integrated Information (IIRC) and its International Integrated Reporting Framework (IIRF) created in 2010, and the Global Reporting Initiative (GRI)'s Sector Supplement for Public Agencies guidelines (Dumay et al., 2010; Farneti & Guthrie, 2009; Montesinos & Brusca, 2019). It must be noted that there exist other guidelines, set by organizations like the European Financial Reporting Advisory Group (EFRAG) or International Financial Reporting Standards Foundation and the like, but studies have shown that the congestion in the number of guidelines has led to a need for political institutions to lay down official guidelines as there appears to be no harmonization in the different guidelines as each organization mainly acts to maintain influence and relevance (Afolabi et al., 2022). For the scope of this research, the pre-dominant IIRF as a basic framework for integrated reporting and GRI guidelines as a basis for sustainability reporting will be taken into consideration due to their breadth of influence especially in the European sphere as the GRI worked with EFRAG to create the aforementioned Directive (Manes-Rossi et al., 2020; Ogata et al., 2018). The details of the two guidelines are seen in the Appendix.

*International Integrated Reporting Framework*

The basis of the IIRF lies in uncovering how an organization creates value and the framework lays this out with three main components: the fundamental concepts that guide the underlying motivation of the report, the guiding principles that specify qualitative points to the report, and the content elements that detail what kind of content fulfills the IIRF. Furthermore, it looks to the governance, strategies, and performance in relation to six "capitals", or "stocks of value that are increased, decreased or transformed through the activities and outputs of the organization" (IRFS, 2021). Although the IIRF was created mainly for private-sector use, the organization says public-sector and non-profit organizations can also use this framework accordingly (IRFS, 2021).

*Global Reporting Initiative*

The GRI specifies sustainability reporting based on economic, social, and environmental performance. They recommend creating reports based on two principles: defining the content and defining the quality (Boiral et al., 2019). In terms of content, the GRI recommends consistency between the information disclosed and the publishing organization. Under this, they recommend further detailed principles such as stakeholder inclusiveness, and sustainability context. In terms of quality, the presentation and quality of information that is presented are prioritized; although it does not specify how the visualization should be conducted, some of their recommending principles include balance, comparability, and reliability " (GRI, 2022). If organizations are not able to comply with the requirements of the GRI Standards, it is able to report "in accordance with the GRI Standards" or "in reference to the GRI Standards" (GRI, 2022, p. 18) due to its modular nature.

*Academic Critique of the Guidelines*

The academic critique of both the IIRC/IIRF and the GRI is abundant. For the IIRC and IIRF, one researcher calls it a "failure" as the core value that drives the IIRC is "value for investors" thereby ignoring the value for society which then furthers and justifies the ignorance of societal effects by the organization (Flower, 2015). Others echo this, saying that the IIRC is focused on a "business case for sustainability rather than the sustainability case for business" (Thomson, 2015).

Research of corporate reports utilizing the GRI has shown the dubiousness of its ability to increase accountability, showcase organizational capacity, and mandate complete reporting of negative events (Afolabi et al., 2022). This is in part due to the GRI's genericness in its guidelines for reporting and therefore does not address specific different

types of organizations (Manes-Rossi et al., 2020). In addition, due to an organization's ability to pick and choose which GRI standard to comply with, the use of GRI by public sector organizations is fragmented with only 3% of the total number of GRI-compliant reports published in 2018 coming from public institutions (Cavalcante et al., 2019; Farneti & Guthrie, 2009; Lodhia et al., 2012), and the reportage between the three dimensions of economy, social and environmental is unequal (Fusco & Ricci, 2019) thereby not giving a complete picture of reality. Furthermore, reports utilizing GRI standards often supplement it with other frameworks, alluding to the insufficiency of the GRI standards (Dumay et al., 2010).

In addition, the GRI is said to be too managerial in that public sector organizations will use the GRIs to simply feed into the managerialist way of evaluating sustainability without resulting in actual sustainability, which Olson et al. refer to as an "evaluator trap" (Dumay et al., 2010; Fusco & Ricci, 2019; 2001, p. 27). Dumay et al. also point to the pointless "accountingization", or numericizing environmental concerns, that do not impact sustainability that may cause "government-sanctioned 'greenwashing' or the development of a platitude of inwardly focused sustainability reports and practices" (2010, p. 544).

The VLRs add to the growing list of the types of reports as it is not exactly SR nor IR or a government report. Although they are formally labeled as "reviews", their definition points to a multi-purpose report to track SDG progress, and this document is to be created by LRGs (United Cities and Local Governments & UN-Habitat, 2021). One would assume that a government report would fit, but a government report traditionally encompasses business reporting such as budgets, service reporting, and infrastructure reporting (Bora et al., 2021), and thus is a mismatch for the VLRs.

# 3 Conceptual Framework

Based on the aforementioned research on performance, performance indicators, and public reports, the section aims to propose a conceptual framework for the analysis of the SDG indicators in the VLRs. The conceptual framework is largely based on the five-step performance measurement process supported by "quality assurance", as described in Section 2.2 and is summarized in Figure 5. The performance measurement process was taken as the basis of the conceptual framework for this research as it encompasses factors that are highly relevant in answering the research questions and is a pre-established way of thinking.

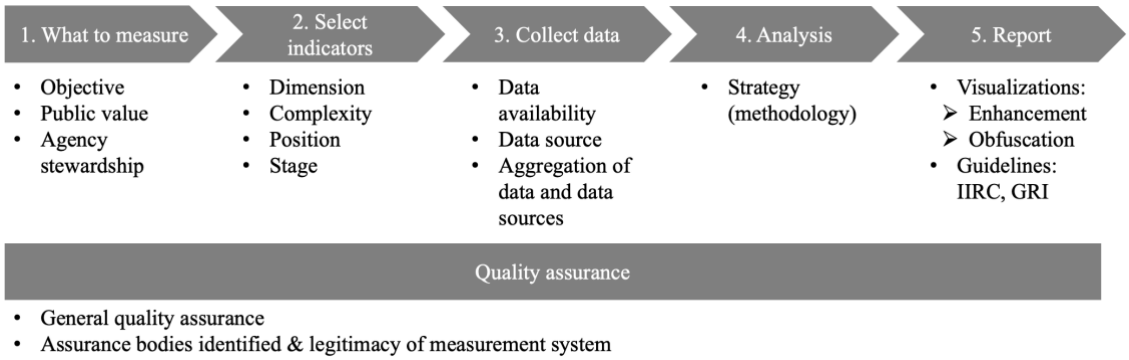



**Figure 5 Conceptual framework for performance measurement assessment**

## 3.1 Description of Proposed Framework

The five steps as previously mentioned are: 1. prioritizing what to measure, 2. selecting indicators, 3. collecting data, 4. analyzing, and 5. reporting. Each step is composed of numerous qualities, taken from various research. These five steps are supported by "quality assurance".

1. *What to measure*

The first step of the process encompasses objective, public value, and agency stewardship. By assessing public value and agency stewardship relationships, this framework attempts to better understand the context of the measurement process within these LRGs. Ultimately, inferences on performance usage including hard and soft use are expected to be found.

2. *Selecting indicators*

The second step of the process looks for the four types of indicators and its sub-types: dimension (single point, time series, cross-section, cross-section, and time series),

complexity (descriptive, ratio, normative, index), position (leading, lagging) and stage (input, output, outcome). By finding the number of certain types and sub-types of indicators used in the VLRs, the advantages and disadvantages of them can be weighed against each other.

3. *Collecting data*

The third step of the process encompasses the three characteristics of data (un)availability, data source, and aggregation of data and data sources.

4. *Analysis*

The fourth step of the process encompasses how the indicators are analyzed, but due to predicted document restrictions, the existence (or lack thereof) of a methodology will be recorded and analyzed.

5. *Report*

The final step of the process encompasses the usage of international guidelines, the usage of visual representations, the kinds of visual representations such as pictographs, line graphs, bar graphs, as well as characteristics of enhancement or obfuscation of these visualizations.

6. *Quality assurance*

This underlying foundation for the performance measurement process is composed of any mentions of qualities that supported the entire process, as well as the identification of assurance bodies and the legitimacy of the measurement system. Qualities can include validity, repeatability, functionality, and others. Legitimacy can include either ownership or general support of organization members, thus the identification of assurance bodies and the legitimacy of the measurement system will be examined in conjunction with one another.

## 3.2 The Soundness of the Proposed Framework

It is said there are four criteria for a sound framework: consistency, testability, empirical accuracy, and simplicity. First, consistency reflects the appropriateness of the framework to the research problem. Second, testability indicates the ability to test hypotheses stemming from the framework in the empirical phase. Third, empirical accuracy refers to the hypotheses created. Lastly, simplicity otherwise known as “parsimony” specifies “the need for sparseness with respect to the number of variables, conditions, and assumptions specified” (S. van Thiel, 2014, p. 56/219). For this kind of inductive research, mainly

consistency, empirical accuracy, and simplicity should be met as testability applies to deductive research. In accordance with these criteria, the proposed conceptual framework was created and edited based on the pre-established performance measurement process with further additions from other researchers in a simple way, which makes it highly relevant to the main research problem as well as empirically accurate.

# 4 Performance Management as Seen in Voluntary Local Reviews

In Section 2, academic research on the related concepts of performance management was explained. In this section, the VLR as an aspect of performance management will be introduced by examining the available gray literature published by organizations such as the United Nations and UN-Habitat, in addition to any published material the LRGs may follow to create their own VLRs. This section aims to understand current gray literature into indicators and presentations of those indicators. In doing so, this section will bridge the academic background introduced in Section 2 and the conceptual framework of Section 3 to further build the foundation for the analysis results in Section 6.

## 4.1 Background of the Voluntary Local Reviews

Aside from the aforementioned definition of the VLRs as "both a reporting tool to assess, monitor and present local achievements in implementing the SDGs, as well as a process to enhance and expand the political and social commitment of a variety of stakeholders to the SDGs, orienting local priorities and development planning" (United Cities and Local Governments & UN-Habitat, 2021), one guide created at the beginning of the VLR history states its objective to "help local organizations to be more efficient, more transparent and more closely connected to their own citizens", mirroring the theories and academic definitions of public reports as stated earlier (UN-Habitat et al., 2019, p. 34). In addition, the specificity of the global indicators helps to establish a baseline that reflects the LRG's capabilities "to the international donor community, or even to potential investors" although ultimately the SDGs should support the LRG's efforts to serve their citizens (UN-Habitat et al., 2019, p. 34). Ultimately, the VLRs can be used internally by LRGs to bring stakeholders together to deepen their understanding of SDG implementation and externally to communicate to other LRGs about local case studies and successes (UN-Habitat et al., 2019, p. 36).

The VLRs are provided with a voluntary guideline by the UNDESA, called the "Global Guiding Elements for Voluntary Local Reviews (VLRs) of SDG Implementation" (UNDESA, 2020). The Global Guiding Elements provide a framework for LRGs to follow based on the "Secretary-General's Voluntary common reporting guidelines for Voluntary National Reviews" (UNDESA, 2020). The underlying guide for the VLR is the 2030 Agenda, and in particular, the VLRs should address five points: be long-term, open, people-centered, based on pre-existing platforms and processes, and evidence-based (UNDESA, 2020). More specifics are given for the proposed structure and content of the report. A summary of the structure is seen in Table 3.

**Table 3 Summary of proposed structure and content for VLR Source: (UNDESA, 2020)**

| No. | Section | Details |
|---|---|---|
| 1 | Opening Statement | Provide opportunity for high-ranking public official to showcase SDG integration in various time horizons, political commitment towards concrete actions in the LRG |
| 2 | Highlights | 1-2 pages main message of report incl. cases that should be shared with other LRGs as it sets the tone of the report, main plans, and challenges of implementation |
| 3 | Introduction | Introduction of LRG, reasons for undergoing the VLR process, overview of SDG implementation incl. progress, actions, strategies, and frameworks |
| 4 | Methodology and process for preparation of the review | Full methodology used for the production of the VLR incl. responsible parties and their assignment process, assignment of priorities of certain SDGs, engagement of local stakeholders, data collection process and its challenges |
| 5 | Policy and enabling environment | |
| 5a | Engagement with the national government on SDG implementation | Main characteristics of the governance system and illustrate the relationship between subnational govt. and national level |
| 5b | Creating ownership of the Sustainable Development Goals and the VLRs | Reflect on raising awareness and publicity of the SDGs, and how they ensured transparency and stakeholder engagement during that process |
| 5c | Incorporation of the Sustainable Development Goals in local and regional frameworks | Map SDG and its targets to local or regional short-, medium-, and long-term plans, strategies and thematic programs and conduct gap analysis to fill in the gaps. Introduce new policies or strategies established to reach the SDGs |
| 5d | Leaving no one behind | Explain how "one of the most powerful tenets of the 2030 Agenda" is translated into local practice |
| 5e | Institutional mechanisms | Reflect on existing mechanisms that serve the LRG in SDG implementation |
| 5f | Structural issues | Reflect on structural issues or barriers such as differences between municipal administrative boundaries and the functional urban area, cross-boundary environmental impacts, decentralization/devolution, etc. |
| 6 | Progress on Goals and Targets | Report on progress and analysis on all 17 Goals to establish baselines, highlight trends, successes, challenges, emerging issues and lessons learned |
| 7 | Means of implementation | Assess existing national and local financial resources, ability of regional or local authorities to raise local resources, their effective allocation and the additional needs for successful implementation in areas such as financing, technology, capacity building, and data |
| 8 | Conclusion and next steps | Outline next steps for enhancing SDG implementation, as part of the Decade of Action and Delivery for sustainable development |
| 9 | Annex | Include statistical annex |

## 4.2 Background of the Indicators in Voluntary Local Reviews

A 320-page report provided by the European Commission's Joint Research Center titled "European Handbook for SDG Voluntary Local Reviews – 2022 Edition" is the latest guideline for European LRGs that gives advice on methods for creating indicators in order to monitor the SDGs appropriately within each local context (Siragusa et al., 2022). Updated from the original 2020 version, the guideline includes 72 indicators tailored to the European context for the 17 SDGs, and steps for how the LRGs can act towards implementing those indicators. The ultimate aim of providing this guideline is to create a baseline of indicators that allow for European-wide comparison. The guideline also provides a look into the evolution of indicators as presented in the numerous VLRs, categorized into three "approaches": using global indicators, using local indicators regardless of SDG target alignment, and using frameworks deriving from either local,

national, or international levels (Siragusa et al., 2022, p. 222). However, the break-out of these approaches means the incomparability of performance assessment between the LRGs even within the same European context. Echoing the aforementioned literature review, this guideline also emphasizes several characteristics of indicators in order to create a VLR that is "considered as an exercise in accountability and transparency" including clear definitions of the said indicator, the methodology for the computation of the indicator, the replicability, the data source, and the observation of "minimum statistical principles" (Siragusa et al., 2022, p. 223).

In preparing for the creation of indicators and the subsequent data collection and analysis, the guideline also points out the inevitable hurdles that LRGs may face, such as disaggregation of data sources, accessibility of data, calculation and analytics in processing the data, and visualizing the findings (Siragusa et al., 2022, p. 223) which reflects the proposed conceptual framework in Section 3.1.

## 4.3 Summary of Pre-Existing Voluntary Local Reviews Analysis

The UN-Habitat is leading the research on the analysis of VLRs; they have published two guidelines as part of the "VLRs Series", a knowledge-sharing initiative (UN-Habitat, n.a.). Titled "Guidelines for Voluntary Local Reviews Volume 1: A Comparative Analysis of Existing VLRs" published in July 2020, and "Guidelines for Voluntary Local Reviews Volume 2: Towards a New Generation of VLRs: Exploring the local-national link" published in July 2021, the contents of these reports will be summarized below.

### 4.3.1 Guidelines for Voluntary Local Reviews Volume 1: A Comparative Analysis of Existing VLRs

As the first VLR analysis, this 51-page report focused on understanding what a VLR is by comparing 39 VLRs around the world based on four variables: the agency driving the VLR, the institutional locus, the technical content, and the data and indicators. For this summary, the technical content and the data and indicators section will be examined in relation to the scope of this research. The analysis of the technical content focused largely on the structure of the then-available VLRs, to reflect on its variety as a continuum as some had strictly adhered to guidelines while others were more explorative, and this was pinned to the degree of VLR institutionalization (United Cities and Local Governments & UN-Habitat, 2020, p. 31). Moreover, the data and indicators section covered the subsequent tasks such as data collection and knowledge management and analyzed the use of indicators, data sources, and the involvement of civil society as presented as columns in a table format spanning five pages. The degree of detail in the "use of indicators" column varied widely, from a detailed percentage of available indicators that

could be viewed online, to a more general acknowledgment of the cherry-picking of indicators or the degree of localization. The "local data" column pointed to data collection and data sources, whether data for SDG indicators are collected on a national or local level. Additionally, other than a brief introduction of localized SDG indicators including the European Office of Statistics Eurostat's European indicators, the analysis of the VLRs found three main methods LRGs were undergoing in creating indicators as a result of lack of data: VLRs using the global indicators but "reworking terminology, method, and sources to make them accessible", VLRs trying to find the correlation between the global indicators and available data, and VLRs that serve as a general qualitative assessment (United Cities and Local Governments & UN-Habitat, 2020, p. 43). In conclusion, the "lessons learned" are summed up three-fold: data management challenges such as disaggregation of national statistics, lack of data, and the technical and financial costs of managing the data, localizing indicators through "adaptation, language reformulation… and crowd-sourced indicators and criteria", and prominent international actors such as the UN-Habitat and the EU that help overcome data management challenges (United Cities and Local Governments & UN-Habitat, 2020).

### 4.3.2 Guidelines for Voluntary Local Reviews Volume 2: Towards a New Generation of VLRs: Exploring the Local-National Link

The second guideline, published a year after the first one, focused on the link between the VNRs and the VLRs. For this summary, one area of focus – the key dimensions in making the linkage work – will be examined, and more specifically, the use of indicators to assess SDG implementation. Acknowledging the controversy of the SDG indicators, the analysis found LRGs struggling with data collection and data measurement, and even the global indicators an obstacle for LRGs to create their own VLRs. In terms of indicator creation, it found participatory and crowdsourcing methods used successfully in cities such as Los Angeles and São Paulo, and in terms of overcoming data management challenges, it pointed to cities like Bogotá in Colombia using those challenges to bridge national and local governments in order to come up with multi-level solutions. Specific analysis into indicators was not conducted within this analysis, assumably due to the scope of this report as pinning VNRs and VLRs to each other. However, one aspect the analysis did touch upon that did not come up in the first guideline was the visualization of the reports, although it was only to the extent of a few sentences. It mentioned that the city of La Paz in Bolivia had "made extensive use of spatial visualization both in the way it showed and communicated data and the way it collected it and designed its indicators", as seen in Figure 6 which depicts income-based poverty and unfulfilled basic needs-based poverty within the different districts of the city (Gobierno Autónomo Municipal de La Paz, 2018, p. 23; United Cities and Local Governments & UN-Habitat, 2021, p. 39).

In using spatial visualization, comprehension of specific challenges within each district is underlined especially as each district has “inherent differences” administratively.

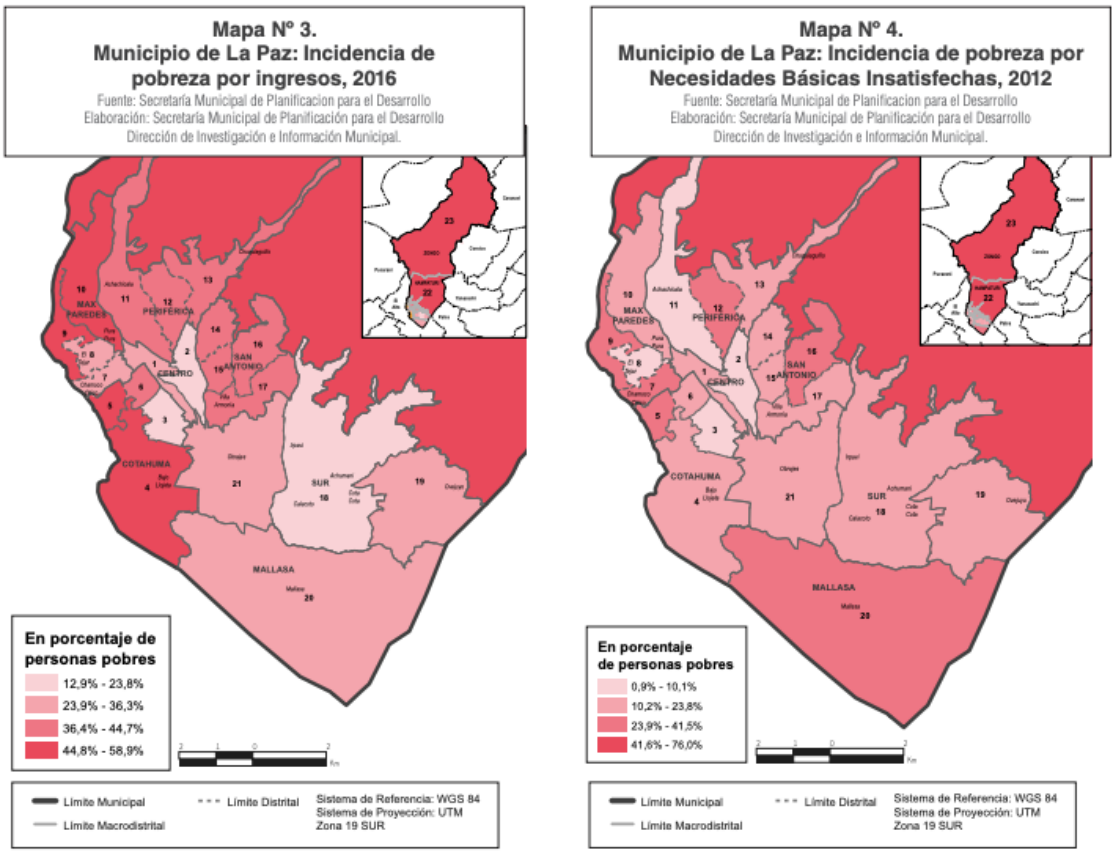



**Figure 6 Sample of spatial visualization from La Paz, Bolivia's VLR (Gobierno Autónomo Municipal de La Paz, 2018, p. 23)**

# 5 Research Approach and Methodology

The research aim was to conduct fundamental research to obtain scientific knowledge, as opposed to conducting applied research in order to solve problems, towards the nature of and presentation of indicators in local SDG plans of European local and regional governments (S. van Thiel, 2014). As local performance tracking of SDG goals is still young and the specific reporting style has only been in place since 2018 (UN-Habitat, 2018), there were observable gaps in academic research thus this research took a descriptive inductive approach.

A framework was then formulated by combining various academically based theories and EU-approved industrial guidelines to increase reliability and validity. This framework would then structure the mixed method case study of all available VLRs in English submitted and published by EU cities. A mix of quantitative methods, mainly descriptive statistics, and qualitative methods mainly content analysis were used. "Case" refers to the report thus delineating the scope of the case study to the observable information of the report. Applying the main research method of content analysis, this research hoped to see general observations as well as similarities and differences between the multiple cases.

This research aimed to achieve reliability through the accuracy and consistency of the variables measured (S. van Thiel, 2014). Accuracy will be attained through the usage of the coding scheme created based on the conceptual framework. Consistency, based on repeatability, will be maintained through the careful documentation of the research steps.

## 5.1 Case Selection and Data Collection

The method of choosing the sampling of the primary material data source uses purposive sampling, based on non-probability sampling. There are currently no worldwide standards for performance indicators and its presentation of the SDGs on a local context. However, one available and widely accepted format is the Voluntary Local Reviews. Therefore, the research decided on the purposive sampling of the VLRs of European Union cities, available on the website of the Secretariat of the United Nations (https://sdgs.un.org/topics/voluntary-local-reviews), and was accessible to the public as of March, 2023. The website listed VLRs of all participating LRG, with the functionality of filtering with individual countries, and provided report for each year available. The researcher chose filtered through each of the 27 EU member states, and manually listed the available LRGs. In the end, there was a total of 24 VLRs originating from nine member states. Some of the LRGs had published VLRs in the past, but to delineate the time horizon of the research, the research chose a cross-sectional time horizon and observed one year per LRG. To ensure the research results would be relevant, the newest

VLR per LRG was chosen. Thus, the report's published year ranged from 2016 to 2022. (Although the "first" VLR was published in 2018 (UN-Habitat, 2018), a report from 2016 was published on the aforementioned website as a VLR so this report was also considered a VLR). Most of the VLRs are available in the English language except for one for the region of Wallonia in Belgium which was only available in the French language. As the researcher is not proficient enough in French, this report was omitted from the case selection, bringing the total number of cases to 23.

To ensure reliability and validity of the data collection, there are three key steps that were identified: the context, quality, and the producer of the data (S. van Thiel, 2014). The context of the VLRs was defined as the LRG, country, year, and language of publication as shown in Table 4. The producer of the data, in this case interpreted to be the editor of the VLRs, are also listed in the same table; if the editor of the VLR was not specified on the VLR, a "not available" is stated. Lastly, the quality of the data interpreted to be synonymous with the main research question of this research, thus will be answered in later sections. Having these details secures the reliability and validity of the data collection.

**Table 4 Data Sources**

| # | Country | Local or regional government | Year | Editor of VLR |
|---|---|---|---|---|
| 1 | Belgium | City of Ghent | 2021 | Ghent Stad |
| 2 | Denmark | Municipality of Gladsaxe | 2022 | Gladsaxe Municipality Strategy, Communication and HR |
| 3 | Finland | City of Tampere | 2022 | *Not available* |
| 4 | Finland | City of Helsinki | 2021 | City of Helsinki, publications of the Central Administration |
| 5 | Finland | City of Vantaa | 2021 | *Not available* |
| 6 | Finland | City of Espoo | 2020 | City of Espoo |
| 7 | Finland | City of Turku | 2020 | City of Turku. Central Administration – Project Development Unit |
| 8 | Germany | City of Bonn | 2022 | Service Agency Communities in One World (SKEW) of Engagement Global gGmbH, Sustainability Network North Rhine-Westphalia e.V, German Council for Sustainable Development |
| 9 | Germany | City of Düsseldorf | 2022 | Service Agency Communities in One World (SKEW) of Engagement Global gGmbH, Sustainability Network North Rhine-Westphalia e.V, German Council for Sustainable Development |
| 10 | Germany | City of Kiel | 2022 | Büro des Stadtpräsidenten \| Office of the City Council President, Internationales und Nachhaltigkeit I International Affairs and Sustainability |
| 11 | Germany | City of Stuttgart | 2021 | State Capital Stuttgart, Administrative Coordination, Communication and International Relations Division, Public Safety, Order and Sport Division |
| 12 | Germany | City of Hanover | 2020 | The Mayor, Directorate of Economic and Environmental Affairs, Agenda 21 and Sustainability Office |
| 13 | Germany | City of Mannheim | 2019 | Department Democracy and Strategy |
| 14 | Germany | State of North Rhine-Westphalia | 2016 | Ministry for Climate Protection, Environment, Agriculture, Conservation and Consumer Protection |
| 15 | Greece | City of Skiathos | 2020 | PlanBe P.C., ANAVATHMISI S.A., the University of the Aegean |

| # | Country | Local or regional government | Year | Editor of VLR |
|---|---|---|---|---|
| 16 | Italy | Region of Lombardy | 2022 | DG Ambiente e Clima, Fondazione Lombardia per l'Ambiente, Presidenza – Programmazione |
| 17 | The Netherlands | Amsterdam | 2022 | The Municipality of Amsterdam; Amsterdam University of Applied Sciences Centre of Expertise for Economic Transformation and Centre of Expertise for Urban Governance and Social Innovation; AMS Institute |
| 18 | Spain | Barcelona | 2021 | Office of the Commissioner for 2030 Agenda |
| 19 | Spain | Basque Country Government | 2021 | *Not available* |
| 20 | Spain | The Generalitat Valencia | 2016 | Directorate General for Cooperation and Solidarity of the Generalitat of Valencia and the ART Initiative of the United Nations Development Programme |
| 21 | Sweden | City of Stockholm | 2021 | Stadsledningskontoret |
| 22 | Sweden | City of Malmö | 2021 | City of Malmö |
| 23 | Sweden | City of Helsingborg | 2021 | Contact Center |

## 5.2 Coding Scheme

Guided by the framework, an inductive approach was taken. The coding scheme based on the conceptual framework is summarized in Table 5, and the VLRs were coded using the software ATLAS.ti and the results were compiled in Microsoft Excel. Table 5 summarizes how the coding was conducted with the detailed descriptions.

**Table 5 Summary of codes table**

| Phase | Category | Sub-category | Description |
|---|---|---|---|
| 1. Priorities | Objective | To learn | Where stated in the VLR with the keyword "objective" and "learn". |
| | | To give account | Where stated in the VLR with the keyword "objective" with other keywords such as "accountability", "responsible". |
| | | To steer and control | Where stated in the VLR with the keyword "objective" and other keywords such as "to control" or "to steer". |
| | Public value | - | |
| | Agency stewardship | - | Where any relationship between the city and other levels of government or public administration was mentioned. |
| 2. Indicators | Dimension | Single point | Measured at a single point in time, or if there is no specification of time comparison. |
| | | Time series | Comparison of one variable over time. |
| | | Cross-section | Comparison of different variables in one defined time. |
| | | Time series and cross-section | Comparison of different variables over time. |
| | Complexity | Description | A description (e.g. "reduced" "increased") without any specific quantified goal. |
| | | Ratio | Comparison between two different variables. A proportion is not a ratio. |
| | | Normative | A description (e.g. "reduced" "increased") with a specific quantified goal or target. |
| | | Composite | Indexes. |
| | Position | Leading | Considering the reliance on context, if there is explicit mention within the text that the indicator in question will be used for forecasting, or if this indicator will lead to something. |
| | | Lagging | Considering the reliance on context, if there is explicit mention within the text that the indicator in question can be considered a result of the past. |
| | Stage | Input | Measurements related directly to the activity conducted (human resources, budgets, materials, etc.). |
| | | Output | Direct, tangible results of a program. |

| Phase | Category | Sub-category | Description |
|---|---|---|---|
| | | Outcome | Measures long-term impact of a program, or value of the program. |
| 3. Data collection | Data availability | - | Where stated in the VLR as “unavailable” or variations of unavailability. |
| | Data source | - | Where stated in the VLR as “source” for indicators. Sources were classified as “internal” when the data came from within the city, if not they were classified as “external” including data from national and sub-regional levels. Whether an organization was internal or external to the city administration was determined via Google search of the organization. “Source” but for non-indicator content such as programs were excluded. |
| | Aggregation of data and data sources | - | Where stated in the VLR with the keyword “aggregation” or “disaggregation”. |
| | Sampling of data and data sources | - | Where stated in the VLR with the keyword “sampling” or “sample”. |
| 4. Analysis | Methodology | - | Where stated in the VLR with the keyword “methodology” or context surrounding how the VLR was created. |
| 5. Report | Guidelines | IIRC | The mention of IIRC or International Integrated Reporting Framework. |
| | | GRI | The mention of GRI or Global Reporting Initiative. |
| | Visual representations | - | If there are any visual representations, “yes”, if not “no”. |
| | Table | Text | If the tables are text-only. |
| | Pictographs | - | If there are pictographs. |
| | Graph type | Line | If there any line graphs. |
| | | Line plus bar | If there are any line plus bar graphs. |
| | | Horizontal bar | If there are any horizontal bar graphs. |
| | | Vertical bar | If there are any vertical bar graphs. |
| | | Pie | If there are any pie charts. |
| | | Spider | If there are any spider charts. |
| | | Map | If there are any maps/spatial visualizations. |
| | Enhancement | - | Measuring the number of pixels and calculating its accuracy, |
| | Obfuscation | - | Examining graph axis. |
| 6. Quality assurance | General quality assurance | - | Where stated in the VLR with the keyword “validity”, “functionality”, or “repeatability”. |
| | Assurance bodies identified & legitimacy of measurement system | - | Where stated in the VLR of relevant stakeholders. |

As stated above, some parts of this analysis require the provision of context, such as why this indicator came to be, or what the LRG hopes to observe with this indicator. However, one of the aims of this research is to understand how the LRG utilizes the VLR as a vehicle for communicating indicators, thus disparity in the detailedness of the indicators is foreseen.

In employing this coding approach, it is imperative to go beyond relying solely on this predetermined list. Researchers must remain cautious about the potential bias of seeking code confirmation, even if it does not accurately capture the underlying reality (Saldaña, 2009). Consequently, this study acknowledged this limitation and adopted a critical stance, reflecting on the relevance and applicability of the codes, while allowing for necessary adjustments to align with the complexities of reality. Ultimately, the analysis integrated documents that offered valuable insights to enrich the research, particularly those relevant to the contextual background of the case.

## 5.3 Data Analysis

The data that resulted from the coding scheme underwent content analysis using both quantitative and qualitative methods. Having this kind of mixed method approach allows for complementarity of the results that otherwise would not have been possible using one method (Kansteiner & König, 2020). For quantitative methods, basic descriptive statistics were used, such as the mean that would reflect the average value of the distribution within a data set, and standard deviation that would reflect the average amount of variability of a data set using Microsoft Excel. For the standard deviation, a low standard deviation would imply the data is centered around the mean which reflects less variability, while a higher standard deviation would imply the data is more spread out. Through these two ways, the general trend of the VLRs as well as the variability of this trend could be uncovered. This quantitative method was complemented by qualitative methods that were guided by the coding scheme. The qualitative analysis mainly searched for patterns in similarity and differences that allowed for further reflection on pre-existing research.

## 5.4 Limitations of Research Method

There are various limitations to this research. In terms of the reliability and validity of this research, two main limitations arise: the researcher's skills and biases, and the measurement instruments of the framework. Although the research is conducted empirically and objectively, the researcher is "only human" and may make mistakes during the course of the research. Additionally, there may be errors in the measurement instrument of the framework as well as later analysis of the results as crafting such framework and deducing analysis requires technique.

Furthermore, there is only one type of data source for the basis of this research: the VLRs. As explained earlier with the nature of reports and indicators, the VLRs may only reflect socially desirable answers. To prevent this in a general scenario, triangulation by means of examining different data sources would have been conducted, such as conducting interviews with parties responsible for creating the VLRs and the indicators. However, due to the nature of the sampling selection process and time constraints, this triangulation of data will not be conducted within the scope of this research. Future research avenues could do so by doing this study with a mixed methods design and incorporating interviews with officials responsible for publishing the VLRs.

# 6 Results

This section will present the results of the document analysis of the 23 VLRs. First, basic observations regarding the VLRs and indicators will be presented. Then, a comparison between the cities of the same countries will be conducted. Finally, the results following the conceptual framework will be presented. In Section 7, the implications of these results will be discussed. Any lists of the local and regional governments presented in this section are in no specific order and do not correlate to anything.

## 6.1 Observations from the VLRs and the presentation of indicators

With the 23 VLRs from nine EU countries, the average VLR was 95 pages with the maximum being 236 pages and the minimum being 40. In addition, there were an average of 66 indicators listed on these VLRs, with the maximum being 203 and the minimum being 28. Three VLR reports (City of Ghent, City of Kiel, and the Generalitat Valencia) did not specify any indicators thus the total number of indicators for these VLRs was counted as zero. Additionally, there were multiple VLRs that listed the total number of indicators but only gave textual explanations for a limited number. As the aforementioned conceptual framework considers context, indicators with explanations were considered. The overall summary can be seen in Figure 7 Summary of page and indicator numbers.

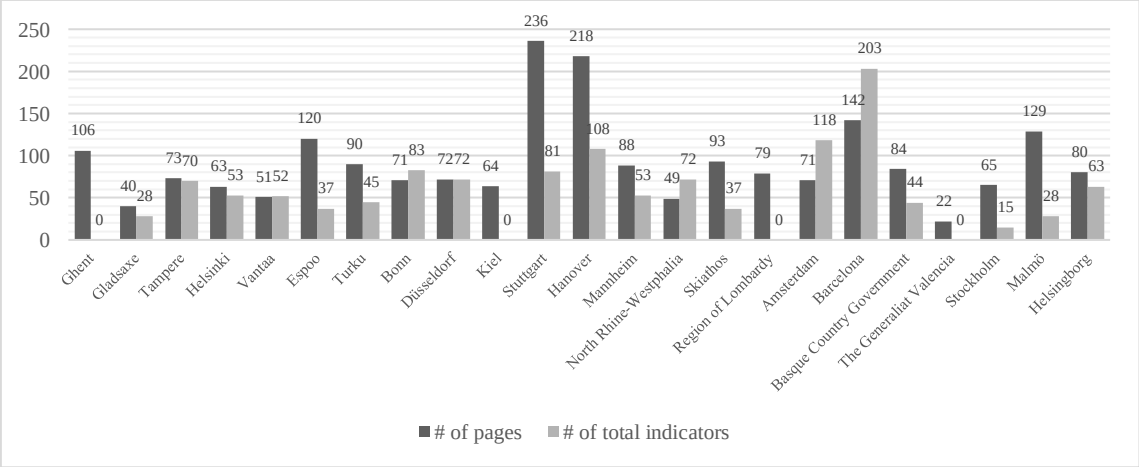



**Figure 7 Summary of page and indicator numbers**

Upon examining all of the VLRs, two general observations were noted about the nature of the VLR and the localization of the SDGs: 1. If the progress of the SDGs was reported based on the 17 goals or local dimensions 2. If local initiatives were mapped out to the SDGs and to what extent (goal-based, or target-based).

The SDGs were introduced in three ways: 1. Individual 17 SDG goals, 2. Grouped together to form local dimensions that parallel local strategies, and 3. none at all. 12 VLRs

used the SDGs as a guide with either all of the 17 goals or the HLPF-designated priority goals to outline progress and within these 12, two focused solely on the HLPF-designated priority goals of that year. On the other hand, 10 used their own dimensions such as the City of Hanover's five dimensions of "environmental, economic, social, cultural, and good governance" (City of Hanover, 2020). Lastly, Generalitat Valencia was the only VLR to not introduce the SDGs as individual goals or through dimensions. The details can be seen in Appendix 6.

All of the VLRs except for the Generalitat Valencia's report "mapped" the SDGs to local indicators or initiatives – there were clear indications of the allocation of local initiatives or indicators to the SDGs or targets. 13 of the reports mapped to goals while eight mapped to the specific targets, as seen in Figure 8 from the City of Düsseldorf. The Generalitat Valencia did not map out specific goals or targets, and Gladsaxe had formed groups of goals based on its six objectives laid out in its "Gladsaxe Strategy 2018-2022" (Gladsaxe Municipality, 2022). The details can be seen in Appendix 6.


**Figure 8 Sample of "mapping out" to SDG targets (City of Düsseldorf, 2022)**

## 6.2 Results based on the conceptual framework

The 23 VLRs were examined based on the aforementioned conceptual framework.

### 6.2.1 What to measure

*Objective*

Between the three objectives for measuring including to give account, to steer and control, and to learn, 19 of the VLRs explicitly stated their reasons for publishing the report and they all had multiple reasons for doing so. (The remaining four VLRs were also the same four VLRs that had not stated they had explicitly created the VLRs for the purpose of creating a review, as mentioned in Section 6.1.) The details can be seen in Appendix 6.

The majority of the reasons pointed to "learning". Within learning, most of the VLRs emphasized the need to understand where they currently are by creating a baseline (City of Amsterdam, 2022; City of Bonn, 2022; City of Ghent, 2021; City of Hanover, 2020; City of Turku, 2020). The second most prominent reason for undertaking the VLR was for steering and controlling, such as the City of Malmö stating, "In this report, the emphasis has been on reviewing the first sub-process – how the City of Malmö's regular

steering and management system steers towards the SDGs" (City of Malmö, 2021, p. 5). In other reports, the VLR was blatantly mentioned as "a strategic tool", or in the City of Ghent's VLR as "to support policies, help prepare them and fuel debates" (City of Ghent, 2021; City of Helsinki, 2021). The least common objective was accountability, with just four VLRs stating so, such as the City of Helsinki saying, "The report seeks to assess the strategy period from the perspective of sustainable development, emphasizing concrete actions and, at the same time, identifying themes for development" (City of Helsinki, 2021). One additional objective that occurred somewhat repeatedly in some of the VLRs was "communicating". Three of the VLRs explicitly stated creating dialogue, bringing stakeholders together for discussions, and relaying the results to citizens as key roles of the VLR (City of Helsingborg, 2021; City of Malmö, 2021; City of Turku, 2020), which hints at the ultimate goal of learning.

*Public value*

Earlier, it was stated that performance can be seen as the realization of public values and thus when examining the VLRs, the underlying values of the cities whether it be about the city in general or about sustainability were scrutinized. eight of the VLRs mentioned "leave no one behind", which is also one of the UN's universal values (United Nations, n.a.-a). Others like the City of Gladsaxe mentioned, "Gladsaxe Municipality must, now and in the future, help our citizens to live independent, active, healthy, and responsible lives, for the greater happiness of the individual and the benefit of society. This requires a local community in balance, both socially and environmentally. It also requires a well-functioning municipality in economic balance" (Gladsaxe Municipality, 2022, p. 6)

*Influences of agency stewardship*

Differing country structures may affect the degree of SDG localization, thus any statements related to country structures were analyzed. There were certainly hints of an agency relationship, such as the Gladsaxe Municipality that stated, "At the request of the Ministry of Finance, Gladsaxe became the first Danish municipality to publish a Voluntary Local Review" (Gladsaxe Municipality, 2022, p. 11). Other hints could be seen in explanations of specific indicators, such as the City of Helsinki stating, "Reducing emissions from traffic requires further efforts and closer cooperation with the state" (City of Helsinki, 2021, p. 19), and the City of Vantaa excluding SDG number 17 because "its sub-objectives were interpreted as state-level work" (City of Vantaa, 2021, p. 8). Additionally, the cities of Bonn and Düsseldorf exhibited an agency relationship with their state and federal government by explaining that the VLR project was developed by the State of North Rhine-Westphalia under the "Global Sustainable Municipality in North Rhine-Westphalia" project and the report was structured under the "German Reporting

Frame Sustainable Municipality" by the German Council for Sustainable Development, a key advisory board to the Federal Government for sustainable development-related topics (City of Bonn, 2022; City of Düsseldorf, 2022; Rat für Nachhaltige Entwicklung, n.a.). Furthermore, the Region of Lombardy strictly outlined the country and region structure with a visual representation delineating scope, as seen in Figure 9.


**Figure 9 Representation of the Lombardy Region Structure as Seen in the VLR (Regione Lombardia, 2022, p. 10)**

### 6.2.2 Indicator selection from the VLR perspective

By explaining the indicator selection stage within the four parts of dimension, complexity, position, and stage and their subsequent sub-categories, this section will allow for further detailed analysis of what kind of indicator is being used currently. In this section, 20 VLRs out of the 23 total VLRs were examined as three of the VLRs did not mention any explicit indicators (City of Kiel, the Generalitat Valencia, City of Ghent). Aside from "position", each indicator usually takes on one characteristic within each of the parts as it has been established that an indicator can be both leading and lagging depending on the context. Each part will be followed by examples of each sub-category. In addition, by dividing the number of indicators allocated to each sub-category by the total number of indicators of each VLR, a percentage could be calculated to better understand the prevalence of each sub-category and so this will also be presented in a table format for each part.

*Dimension*

The "dimension" category is composed of four sub-categories: single-point, time series, cross-section, and cross-section and time series. It was found that more often than not, the indicators would not specify the scope of measurement including the time span in the indicator itself. Therefore, like in Table 6, the researcher included "(with additional context)" to supplement the analysis in order to accurately categorize the indicator.

**Table 6 Example of dimension indicators**

| Dimension | Example indicator (with additional context) | Country | Local or regional government |
|---|---|---|---|
| Single-point | General at risk of poverty rate in the municipality (for 2020) | Finland | City of Tampere |
| Time series | Percentage of home- less people assisted (with trend from last year) | Spain | City of Barcelona |
| Cross-section | Going short of food, medicines, or physician visit because of lack of money (as seen in different cities and ages) | Finland | City of Espoo |
| Cross-section and time series | General at-risk-of-poverty rate (number of people and yearly) | Finland | City of Espoo |

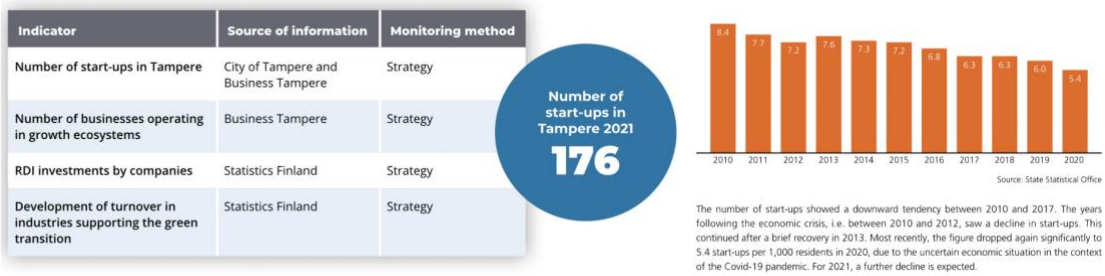
**Table 7 Summary of the dimension sub-category prevalence**

| City | Single-point | Time series | Cross-section | Cross-section and Time series |
|---|---|---|---|---|
| **Total average** | **12%** | **46%** | **5%** | **35%** |
| **Standard deviation** | **26%** | **38%** | **10%** | **35%** |
| Municipality of Gladsaxe | 0% | 43% | 0% | 57% |
| City of Tampere | 100% | 0% | 0% | 0% |
| City of Helsinki | 0% | 43% | 0% | 57% |
| City of Vantaa | 66% | 4% | 25% | 6% |
| City of Espoo | 0% | 19% | 35% | 46% |
| City of Turku | 13% | 73% | 0% | 0% |
| City of Bonn | 0% | 78% | 1% | 19% |
| City of Dusseldorf | 0% | 86% | 1% | 10% |
| City of Stuttgart | 1% | 64% | 3% | 34% |
| City of Hanover | 4% | 16% | 6% | 73% |
| City of Mannheim | 0% | 98% | 0% | 0% |
| State of North Rhine-Westphalia | 0% | 98% | 0% | 0% |
| City of Skiathos | 0% | 84% | 0% | 0% |
| Region of Lombardy | 0% | 0% | 0% | 97% |
| City of Amsterdam | 30% | 11% | 24% | 34% |
| City of Barcelona | 12% | 78% | 0% | 9% |
| Basque Country Government | 0% | 98% | 0% | 2% |
| City of Stockholm | 0% | 13% | 0% | 88% |
| City of Malmo | 7% | 4% | 11% | 74% |
| City of Helsingborg | 0% | 15% | 2% | 90% |

With an average of 46%, most of the indicators were classified as time series with 18 VLRs using them, and then cross-section and time series with 15 VLRs using them.

However, the standard deviation for the time series indicators was also the highest, reflecting a high variability in the percentage of indicators that were classified as such. Single-point indicators were used the least, but the City of Tampere's indicators were all classified as such due to the indicators' lack of details regarding time or coverage. For example, one of their indicators was "the number of start-ups in Tampere" and it stated in the VLR that there were 176 start-ups in 2021 (City of Tampere, 2022). However, the VLR did not go further into whether this was an increase or decrease from previous years, if there was a goal in terms of the number of start-ups, or how this number compared to other cities. On the other hand, the City of Stuttgart had the same indicator but as seen in Figure 10, there is a time-series comparison that allows the reader to better understand what the result of this indicator signifies.

**Figure 10 Single-point versus time-series example from the City of Tampere (2022) on the left and the City of Stuttgart (2021) on the right**




Cross-section indicators were presented largely in two ways: a breakdown of the results into different attributes, or a comparison between different cities. For example, the City of Espoo for SDG 8's "promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all" presented two indicators: population by activity, and state taxable median income (City of Espoo, 2020). For the former, they broke down the population into six attributes: employed, unemployed, student, pensioners, 0-14 years old, and others. For the latter, they took the median income for Espoo as well as five other cities and the national average for a more comprehensive comparison. In the same VLR, they compared Espoo to cities such as Helsinki, Vantaa, Tampere, Turku, and Oulu for their "health barometer morbidity index 2018"; Helsinki and Vantaa both morbidity in their VLRs but the former did not mention this index and the latter broke down its morbidity index into age groups and districts, while Tampere and Turku did not mention morbidity at all (City of Espoo, 2020, p. 224; City of Helsinki, 2021, p. 26; City of Tampere, 2022; City of Turku, 2020; City of Vantaa, 2021, p. 34)

*Complexity*

The complexity category is followed by four sub-categories: description, ratio, normative and composite.

**Table 8 Example of the complexity indicators**

| Complexity | Example indicator (with additional context) | Country | Local or regional governmen |
|---|---|---|---|
| Descriptive | The proportion of the unemployed gaining work within 3 months of concluding business-oriented activation must increase. | Denmark | Municipality of Gladsaxe |
| Ratio | Exercising less than one hour per day 2019 education type girls/boys | Finland | City of Vantaa |
| Normative | The number of jobs in the private sector must increase by 2,000 over 4 years. | Denmark | Municipality of Gladsaxe |
| Composite | Total of human interventions in ecosystems (hemeroby index) | Germany | City of Bonn |

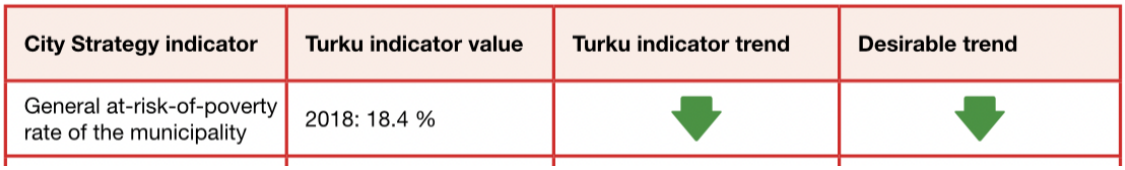
**Table 9 Summary of the complexity sub-category prevalence**

| City | Descriptive | Ratio | Normative | Composite |
|---|---|---|---|---|
| **Total average** | **78%** | **1%** | **13%** | **1%** |
| **Standard deviation** | **23%** | **9%** | **22%** | **2%** |
| Municipality of Gladsaxe | 79% | 0% | 21% | 0% |
| City of Tampere | 96% | 0% | 0% | 3% |
| City of Helsinki | 91% | 6% | 2% | 2% |
| City of Vantaa | 83% | 0% | 11% | 4% |
| City of Espoo | 97% | 0% | 0% | 3% |
| City of Turku | 9% | 2% | 82% | 2% |
| City of Bonn | 63% | 1% | 39% | 1% |
| City of Dusseldorf | 60% | 1% | 41% | 1% |
| City of Stuttgart | 96% | 1% | 1% | 3% |
| City of Hanover | 96% | 0% | 2% | 0% |
| City of Mannheim | 96% | 0% | 2% | 0% |
| State of North Rhine-Westphalia | 61% | 0% | 38% | 0% |
| City of Skiathos | 84% | 0% | 0% | 0% |
| Region of Lombardy | 97% | 0% | 0% | 0% |
| City of Amsterdam | 50% | 2% | 0% | 1% |
| City of Barcelona | 86% | 0% | 0% | 2% |
| Basque Country Government | 91% | 0% | 2% | 5% |
| City of Stockholm | 100% | 0% | 0% | 0% |
| City of Malmo | 59% | 0% | 0% | 0% |
| City of Helsingborg | 64% | 0% | 8% | 0% |

With an average of 78%, most of the indicators in the 20 VLRs were descriptive; they simply described the indicator without any acknowledgment of timelines, goals, or indexes that would have made them normative or composite indicators. Compared to the results from the dimension type, the standard deviation of the prevalence of descriptive

indicators was lower, meaning that a good portion of these indicators are indeed descriptive. Many of the indicators presented the results in terms of percentage (e.g. “Satisfied with his or her life at the moment, % of pupils” (City of Espoo, 2020)), while the usage of ratios was much less. An example of a ratio as an indicator comes from the City of Amsterdam with their indicator “energy ratio” to manage SDG 7, and this ratio was defined as “the percentage of a household income spent on energy” (City of Amsterdam, 2022). Furthermore, normative indicators were not as common as descriptive indicators. The City of Turku’s indicators were mostly normative as they expressed a “desirable trend” in the form of traffic light arrows for each of the indicators as seen in Figure 11 in addition to the actual trend and the indicator value, thus hinting at some sort of goal. Finally, the use of composite indicators was also limited, and when it was used, there appeared to be no specific pattern; indexes were used for measuring education level, gender equality, morbidity, public transport, air quality, soil index, and street safety (Basque Government, 2021; City of Amsterdam, 2022; City of Bonn, 2022; City of Düsseldorf, 2022; City of Espoo, 2020; City of Tampere, 2022; City of Turku, 2020; City of Vantaa, 2021; State Capital Stuttgart, 2021).

**Figure 11 Sample of normative indicator presentation (City of Turku, 2020)**




*Position*

The position category is followed by two sub-categories: leading and lagging. As mentioned in the conceptual framework, categorizing indicators into leading and/or lagging requires context that some of the VLRs did not provide. Therefore, results were only recorded if there were explicit mention of leading or lagging characteristics i.e. what the city hopes to understand from the specific indicator, or what the city assumes is causing the indicator.

**Table 10 Example of the position indicators**

| Position | Example indicator | Context | Country | Local or regional government |
|---|---|---|---|---|
| Leading | Average net rent per square meter | "The topic of housing plays a central role in everyone‘s life and strongly impacts the quality of life." | Germany | City of Bonn |
| Lagging | Average life expectancy | "There is a clear link between life expectancy and education for both men and women. Residents with a lower secondary education alone and a low income have poorer health and the largest number of risk factors for disease. For example, daily smoking is three times as | Sweden | City of Stockholm |

| Position | Example indicator | Context | Country | Local or regional government |
|---|---|---|---|---|
| | | common among people with only a lower secondary education than among those with tertiary education." | | |

**Table 11 Summary of the position sub-category prevalence**

| City | Leading | Lagging |
|---|---|---|
| Municipality of Gladsaxe | - | - |
| City of Tampere | - | - |
| City of Helsinki | - | - |
| City of Vantaa | - | - |
| City of Espoo | - | - |
| City of Turku | - | - |
| City of Bonn | 57% | 3% |
| City of Dusseldorf | 46% | 1% |
| City of Stuttgart | 34% | 11% |
| City of Hanover | 4% | 2% |
| City of Mannheim | - | - |
| State of North Rhine-Westphalia | - | - |
| City of Skiathos | - | - |
| Region of Lombardy | - | - |
| City of Amsterdam | 2% | 3% |
| City of Barcelona | - | - |
| Basque Country Government | - | - |
| City of Stockholm | 50% | 25% |
| City of Malmo | - | 4% |
| City of Helsingborg | 10% | 7% |

According to these classifications, less than half of the VLRs acknowledged how measuring that specific indicator would have potential impacts or causes. Based on the correlation between the percentage of leading and lagging, it was found that once a VLR explains a leading or lagging factor for one of their indicators, they are more likely to explain other indicators, as seen in Appendix 9. Although by definition, indicators can be both leading and lagging, the indicators of the VLRs were usually either or, with a few exceptions. For example, the City of Stuttgart has an indicator "soil index" and they explain the significance of this indicator as follows:

> "*Soil is one of the resources that can hardly be regenerated in human time periods. Therefore, the economic management of local soil resources is a key component of successful concepts for sustainable soil protection. Since classic consumption patterns, such as the construction of detached houses on the outskirts, inevitably drain the resource and land use cannot be balanced effectively, sustainability can only be achieved if a constant, preferably good soil condition (i.e. a defined standard of functional*

*compliance = soil quality) can be guaranteed during a defined period under review. This is only possible if the new use of land is reduced consistently to achieve circular land use management. The focus of soil protection efforts in the State Capital Stuttgart is on the preservation of multifunctional soils, which fulfil the soil functions." (State Capital Stuttgart, 2021, p. 151)*

This implies that the indicator is a lag (or result) of successful economic management, as well as a lead (or forecast) of circular land use management, which then can be used to assess related SDGs such as SDG 6's "Clean Water and Sanitation" (State Capital Stuttgart, 2021).

*Stage*

The stage category is composed of three sub-categories: input, output, and outcome. Although it can also be argued that determining input from output and outcome may require context, for this research, input indicators were classified as characteristics related to budgets, human resources, and materials.

**Table 12 Example of the stage indicators**

| Stage | Example indicator | Country | Local or regional government |
|---|---|---|---|
| Input | Number of training places available per 1,000 applicants for a training place (from October to September of the following year) | Germany | City of Bonn |
| Output | Degree of domestic origin of food purchased by Pirkanmaa Voimia | Finland | City of Tampere |
| Outcome | Proportion of persons at risk of relative poverty, considering the poverty threshold in the autonomous community | Spain | Basque Country |

**Table 13 Summary of stage sub-category prevalence**

| City | Input | Output | Outcome |
|---|---|---|---|
| **Total average** | **1%** | **54%** | **42%** |
| **Standard deviation** | **2%** | **28%** | **29%** |
| Municipality of Gladsaxe | 0% | 36% | 64% |
| City of Tampere | 0% | 53% | 47% |
| City of Helsinki | 0% | 40% | 60% |
| City of Vantaa | 0% | 38% | 60% |
| City of Espoo | 0% | 54% | 46% |
| City of Turku | 0% | 47% | 53% |
| City of Bonn | 4% | 59% | 35% |
| City of Dusseldorf | 1% | 47% | 47% |
| City of Stuttgart | 8% | 29% | 63% |
| City of Hanover | 0% | 96% | 0% |
| City of Mannheim | 0% | 98% | 0% |
| State of North Rhine-Westphalia | 3% | 95% | 0% |

| | | | |
|---|---|---|---|
| City of Skiathos | 0% | 84% | 0% |
| Region of Lombardy | 0% | 97% | 0% |
| City of Amsterdam | 0% | 56% | 44% |
| City of Barcelona | 0% | 57% | 43% |
| Basque Country Government | 0% | 14% | 84% |
| City of Stockholm | 0% | 0% | 100% |
| City of Malmo | 0% | 44% | 44% |
| City of Helsingborg | 0% | 41% | 59% |

With an average of 1% and a standard deviation of 2%, inputs such as human resources or budget were rarely used as indicators. The majority of indicators were either output or outcome. For indicators that were labeled as input, they included things such as "total municipal debt" and "cash surplus/requirement for the long-term fulfillment of tasks" from the City of Stuttgart. It is interesting to note that the former indicator had clear indications it is considered a leading indicator, as it stated that "only with a stable budget situation can the municipality react to problems and undesirable developments" (State Capital Stuttgart, 2021, p. 162), but for the latter, it has clear indications it is considered a lagging indicator with the explanation: "The indicator provides information on the extent to which a municipality is able to make regular payments for current administrative activities from its own resources and without borrowing" (State Capital Stuttgart, 2021, p. 163).

### 6.2.3 Data collection

In this section, the indication of data (un)availability, data sources, and aggregation/disaggregation and sampling of data and data sources will be examined.

*Data (un)availability*

In terms of data (un)availability, only several VLRs mentioned this. There was one VLR from the Lombardy Region that explicitly stated their selection of indicators came from the "availability of data on a local scale" (Regione Lombardia, 2022, p. 13).

There were largely two kinds of (un)availability: not having the data at all or having the data for some years but not for others. An example of the former comes from the City of Turku, where they had several indicators with no data available. One of these indicators was "residents' participation in city development projects"; despite not having the data, they explained the significance of this indicator as a reflection of the residents' experiences in Turku, the Market Square as a meeting place for residents, and the city's investment in digital services for inclusion (City of Turku, 2020, p. 54).

An example of the latter comes from Gladsaxe Municipality; they had one indicator ("The proportion of young citizens (aged 18-29) receiving public benefits must be reduced")

that was measured until 2020 but was unable to collect data for 2021 because of a "transition to a different administrative system". Although the report does not explicitly mention if they will reuse this indicator in the future or updates to the new administrative system, the report states "there will be an additional focus on frequent interviews and on young people participating in job or educational activity" (Gladsaxe Municipality, 2022, p. 20).

*Data sources*

Out of the 23 VLRs, 15 stated data sources for their indicators, results, and any graphs or statistics that they presented. Unlike cities such as Tampere that presented its indicators in a table format and had a dedicated "source of information" column, some VLRs like the Gladsaxe Municipality did not explicitly list data sources next to their indicators but instead would write the data sources in the surrounding text that gave context to the indicators that could be categorized as having an "indirect source" (City of Tampere, 2022, p. 12; Gladsaxe Municipality, 2022, p. 22). Additionally, the City of Espoo would have data sources for certain indicators and their results, but not for others (City of Espoo, 2020, p. 231). The varying degree of data source indication could be summarized into four categories, as seen in Table 14.

**Table 14 Summary of the varying degree of data source indication**

| Categories of data sources | Number of VLRs |
|---|---|
| 1. For every data point, there is a direct source. | 9 |
| 2. For some data points, there is a direct source. | 5 |
| 3. For some data points, there is an indirect source. | 2 |
| 4. There are no data sources. | 7 |

Furthermore, the kinds of data sources varied from international data sources to localized forms of data sources, with a summary seen in Table 15. Within a single VLR, various indicators would use the same source thus only unique indicators were recorded. Data sources varied from pre-published annual reports, websites, maps, universities, and of course statistics offices. They were further categorized into external and internal data sources respectively, with internal data sources classified as such based on whether they came directly from the city administration and all remaining data sources classified as external. The details can be found in Appendix 7.

In terms of general observations, all of the VLRs that indicated data sources had used a mix of internal and external sources, with some using more external than internal, and vice versa. The specificity of the data sources also varied: within the same country, the city of Helsingborg would denote specific ID numbers from its national statistics service Kolada while the city of Stockholm simply wrote "Kolada" without any reference to ID

numbers. The City of Amsterdam from The Netherlands was also very particular about data sources and would include the page numbers from pre-published reports. Furthermore, it was interesting to see that the cities of Düsseldorf and Stuttgart both from Germany both measured gross-domestic-product (GDP) but using different sources: the former took its GDP from the international SDG Portal while the latter made its own calculations from internal sources.

As far as the researcher could tell, there was only one instance in which the data source for an indicator changed over the years; the City of Düsseldorf for the indicator "wastewater treatment: the proportion of wastewater treated by nitrogen and phosphorus elimination" had data available for the years 2010, 2015, and 2018 but for 2010, the data came from the City while the rest came from the SDG Portal (City of Düsseldorf, 2022, p. 54). There was no reason given for the switch, but the results of the indicator for all the years were all the same.

**Table 15 Summary of data sources**

| City | Example data sources | Interna/External classification |
|---|---|---|
| City of Ghent | OECD, City of Ghent Childcare Department | External, Internal |
| Municipality of Gladsaxe | N/A | N/A |
| City of Tampere | Statistics Finland, City of Tampere | External, Internal |
| City of Helsinki | Statistics Finland, Helsinki Welfare Report 2018-2021 | External, Internal |
| City of Vantaa | N/A | N/A |
| City of Espoo | Statistics Finland, espoo.fi | External, Internal |
| City of Turku | National study on passenger traffic, WSP Finland Oy Hyvinvointikompassi, THL | External |
| City of Bonn | N/A | N/A |
| City of Dusseldorf | SDG Portal, State Capital Düsseldorf | External, Internal |
| City of Kiel | SDG Portal, https://www.kiel.de/de/kiel_zukunft/ nachhaltigkeitsziele/hochwertige_bildung.php | External, Internal |
| City of Stuttgart | State Statistics Office of Baden-Württemberg, State Capital Stuttgart | External, Internal |
| City of Hanover | Statistical Office of the Federal State of Lower Saxony Statistical Yearbook 2019, City of Hanover | External, Internal |
| City of Mannheim | N/A | N/A |
| State of North Rhine-Westphalia | N/A | N/A |
| City of Skiathos | N/A | N/A |
| Region of Lombardy | N/A | N/A |
| City of Amsterdam | MRA, Economische Verkenningen Metropoolregio Amsterdam 2021, Staat van de Stad 2021 | External, Internal |
| City of Barcelona | N/A | |
| Basque Country Government | Basque Sociometer Office of Sociological Research Presidency of the Government | Internal |
| The Generalitat Valencia | N/A | |
| City of Stockholm | Sweco, Environment Programme 2020–2023 the City of Stockholm. | External, Internal |
| City of Malmo | Kolada, City of Malmo | External, Internal |
| City of Helsingborg | Kolada, City of Helsingborg | External, Internal |

*Aggregation/disaggregation and sampling of data and data sources*

Regardless of if it was occurring, the aggregation/disaggregation or the sampling of data and data sources were not commonly indicated; only five VLRs stated aggregation/disaggregation, and three VLRs stated sampling. When aggregation was mentioned, the limits it imposed on analysis were also mentioned: “The collection of data disaggregating different population groups is limited in Vantaa, and it is precisely such data that is essential to make the spirit of Agenda 2030 a reality - no one is left behind” (City of Vantaa, 2021, p. 50). Additionally, the City of Kiel mentioned sampling when explaining one of its statistics: “The member organizations surveyed do not reflect the entire range of providers but provide a useful insight into the development of the provision landscape in Kiel” (Landeshauptstadt Kiel, 2022, p. 25).

### 6.2.4 Analysis

In an attempt to find out the analysis methods taken in the VLRs, the methodology of how the VLR was conducted was analyzed for this section. In the end, 14 out of the 23 VLRs (Cities of Vantaa, Espoo, Turku, Bonn, Düsseldorf, Skiathos, Malmö, Amsterdam, Hanover, Municipality of Gladsaxe, Region of Lombardy) had dedicated some sort of methodology section within the report. Seven VLRs (Cities of Barcelona, Ghent, Tampere, Mannheim, Stockholm, Helsingborg, State of North Rhine-Westphalia, Basque Country Government) had no dedicated methodology section, while one VLR (City of Stuttgart) also had no methodology section for the report but instead had a methodology for each of the indicators.

For the reports that did include a methodology, the contents of the methodology widely varied in length and detail. Some defined overarching objectives and aims of the review that guided the report, such as the City of Espoo with nine goals such as the review being “factual, concrete and solution-oriented” (City of Espoo, 2020, p. 16).

Others listed the details of the review process, including when it started, who within the city was responsible for what, and what kind of analysis took place. One way of representation was with a timeline, as seen in Figure 12. Three VLRs (Cities of Helsinki and Kiel, the Generalitat Valencia) had this timeline of the VLR creation and implementation. Another example of the review process comes from the Gladsaxe Municipality which stated, “The process of the review reaches across all sectors within the organization and takes its point of departure in feedback once a year. All areas are asked to report on the progress and comment on deviations from the overall direction” (Gladsaxe Municipality, 2022, p. 8). The City of Espoo had a detailed explanation as well, stating:

*"The review process started at the end of January 2020. In February, the process and the manner in which the review was to be carried out were planned and decided on. The most active review work was carried out in March and April 2020. All phenomena were reviewed at least three times by five different persons. All in all review process with cross checking lasted until late stages of layout phase"* (City of Espoo, 2020, p. 16).

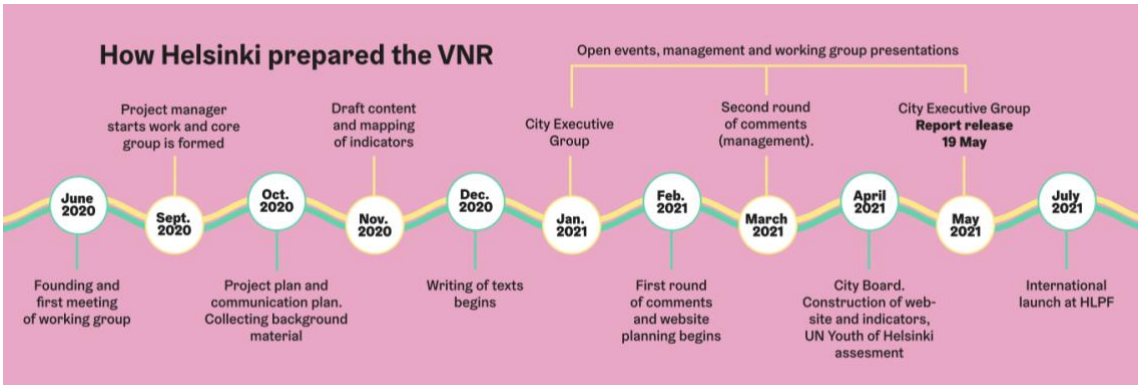



**Figure 12 Sample of methodology representation in timeline format (City of Helsinki, 2021, p. 10)**

Other reports had clear objectives for the indicators and SDGs, such as the Gladsaxe Municipality stating, "The targets define the desired direction and level of progress during the 4-year political term set by the City Council" (Gladsaxe Municipality, 2022, p. 8) and the City of Vantaa stating, "In the selection of goal-specific indicators, emphasis was placed on relevance from Vantaa's point of view and compliance with the UN's sub-goals for sustainable development. The indicators were selected from among existing indicators" (City of Vantaa, 2021, p. 8). Additionally, the City of Amsterdam had chosen specific SDGs due to reasons such as those goals' appearance in other VLRs and the availability of policy documents (City of Amsterdam, 2022, p. 127).

Furthermore, some reports had descriptions for data collection, including whether they conducted qualitative and/or quantitative studies. The cities of Bonn and Düsseldorf, which created the VLRs under a shared framework thus sharing the methodology, stated:

*"A comprehensive stocktaking was conducted at the beginning of the project to gather the information needed for this report. This process consisted of both qualitative and quantitative elements. In terms of the qualitative analysis, all core activities in the city that contribute to sustainable development and represent the status quo were gathered. This procedure included integrated or sectoral strategies and concepts, measures, projects, city council resolutions, specific goals, collaborations*

*and networks, and organizational structures."* (City of Bonn, 2022, p. 13; City of Düsseldorf, 2022, p. 13)

Lastly, one VLR had a section titled "methodology" that stated relevant documents that the VLR was based on including the VLR handbooks and international standards, and four phases that it took for the city to complete the review (Municipality of Skiathos, 2020, p. 7).

Taking these details into account, the stated methodologies of the VLRs could be inductively categorized in five ways: 1. Definition of objectives 2. Process of review 3. Data collection 4. (Role of) indicators 5. Other. A summary of these six can be found in Table 16.

**Table 16 Summary of methodology contents**

| Aspects of VLR methodology | Definition | VLR |
|---|---|---|
| Definition of objectives | Objectives, goals, and aims of the review | Cities of Espoo, Hanover |
| Process of review | Timeline, who did what kind of analysis | Cities of Vantaa, Espoo, Bonn, Düsseldorf, Helsinki, Kiel, Gladsaxe Municipality, Generalitat Valencia |
| Data collection | Specification of quantitative and/or qualitative analysis | Cities of Espoo, Bonn, Düsseldorf, Malmö, Gladsaxe Municipality |
| (Role of) Indicators | Objectives, goals, aims, and derivation of indicators | Cities of Vantaa, Bonn, Düsseldorf, Hanover, Gladsaxe Municipality |
| Other | Other | City of Skiathos |

### 6.2.5 Report

Based on the conceptual framework, data was collected on seven main points about reports: if it had followed the international guidelines, if there were any visual representations, if there were tables, pictographs, differing graph types, and if there were any signs of enhancement or obfuscation as defined in the conceptual framework based on the literature review.

*International guidelines*

Out of the 23 VLRs, only one (City of Skiathos) mentioned the international guidelines. Skiathos mentioned the Global Reporting Initiative as one of the documents that the city had looked to for international standards (Municipality of Skiathos, 2020, p. 7). Interestingly, there seemed to be no other mention of any of the VLRs using guidelines for the creation of these reports, although some did mention guidelines for the indicators such as cities of Bonn, Düsseldorf, and Amsterdam that had based their VLRs on "the Guidelines for Voluntary Local Reviews by UCLG and UN-Habitat, the European Handbook for SDG Voluntary Local Reviews by the European Commission's Joint

Research Centre, and the Global Guiding Elements for Voluntary Local Reviews of SDG implementation by UNDESA" (City of Amsterdam, 2022; City of Bonn, 2022; City of Düsseldorf, 2022).

Additionally, in the State of North-Rhine Westphalia, the "Global Sustainable Municipality in North Rhine-Westphalia" project was implemented in which the VLR itself was structured with the "German Reporting Frame Sustainable Municipality" published in 2021 as a country-wide guideline of sorts (City of Bonn, 2022; City of Düsseldorf, 2022). Bonn and Düsseldorf both created the VLR under this project and structure, and it was also stated that the City of Münster, City of Dortmund, and the City of Arnsberg had already created VLRs under this initiative but these three reports were not seen on the UN's official VLR page (City of Bonn, 2022). Regardless, the only commonalities of the Bonn and Düsseldorf VLR seemed to rest on the surface: the design of the VLR in terms of colors and font was the same as well as the overall structure with the same headings. They also shared the same text in certain parts and for some of the indicators, but the commonalities stopped there. In terms of structure and "rules" of the VLRs, the mention of the data sources was present in Düsseldorf but not for Bonn, and Bonn showed progress via the use of arrows next to each indicator while Düsseldorf had no progress visualizations as seen in Figure 13 with Düsseldorf to the right and Bonn to the left (City of Bonn, 2022; City of Düsseldorf, 2022).


**Figure 13 Example of two German VLR representation (City of Bonn, 2022; City of Düsseldorf, 2022)**

*Visual representations*

There were 17 VLRs (Cities of Ghent, Helsinki, Vantaa, Espoo, Turku, Bonn, Düsseldorf, Kiel, Stuttgart, Hanover, Amsterdam, Skiathos, Stockholm, Malmö, Helsingborg, Barcelona, and the Region of Lombardy) that had used some sort of visual representation, while the remaining reports only had textual content. Out of the 15, the most used form of representation was the vertical bar graph and the line graph with 12 VLRs, with the details below. The VLRs with the most types of visualizations was the City of Espoo which used all the forms of visualizations except for spider graphs. In second was using

five types of visualizations, which the Cities of Helsinki, Amsterdam, Ghent, and Hanover had done. In total, the average number of visualizations used in each VLR was 2.3 visualizations. The details can be seen in Appendix 8

**Table 17 Summary of types of visualizations**

| Type of visualization | Number of VLRs (%) | VLR |
|---|---|---|
| No visualization | 6 (26%) | Cities of Tampere, Turku, Mannheim & State of North Rhine-Westphalia & Basque Country Government & the Generaliat Valencia |
| Line | 12 (52%) | Cities of Ghent, Helsinki, Espoo, Bonn, Düsseldorf, Stuttgart, Hanover, Amsterdam, Stockholm, Malmo, Helsingborg, & Region of Lombardy |
| Line plus bar | 5 (21%) | Cities of Ghent, Espoo, Hanover, Malmo, Barcelona |
| Horizontal bar graph | 4 (17%) | Cities of Espoo, Bonn, Düsseldorf, Vantaa |
| Vertical bar graph | 12 (52%) | Cities of Espoo, Bonn, Düsseldorf, Hanover, Malmo, Stuttgart, Amsterdam, Stockholm, Helsingborg, Ghent, Helsinki & Region of Lombardy |
| Pie chart | 6 (26%) | Cities of Espoo, Hanover, Stuttgart, Amsterdam, Helsingborg, Vantaa |
| Spider chart | 2 (8%) | City of Helsinki, Region of Lombardy |
| Map/spatial visualization | 6 (26%) | Cities of Helsinki, Amsterdam, Malmo, Espoo, Vantaa, Ghent |
| Pictograph | 5 (21%) | Cities of Espoo, Turku, Kiel, Skiathos, Amsterdam |

Additionally, eight VLRs utilized pictographs in a blend of two ways: arrows and traffic lights. Traffic lights use a three-color palette where green signals good/positive, yellow signals neutral, and red signals bad/negative. Three VLRs used arrows, three used traffic lights, and three used the blend of arrows with traffic lights.

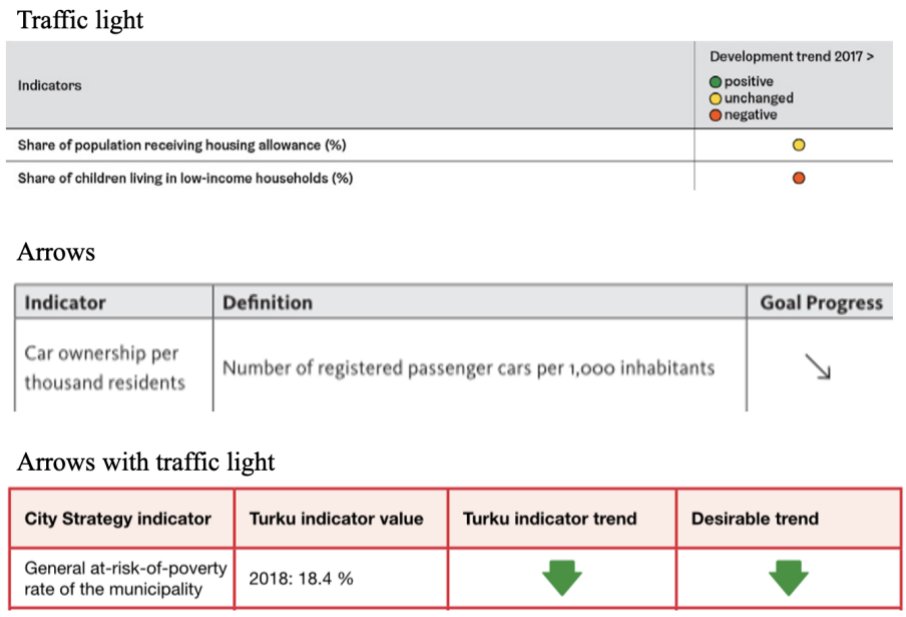



**Figure 14 Samples of pictographs from top to bottom: Cities of Helsinki, Mannheim, Turku (City of Helsinki, 2021, p. 59; City of Mannheim, 2019, p. 63; City of Turku, 2020, p. 14)**

**Table 18 Summary of pictograph usage**

| Type of visualization | Number of VLRs (%) | VLR |
|---|---|---|
| Arrows | 3 (13%) | Cities of Mannheim, Helsingborg, and Basque Country Government |
| Traffic lights | 3 (13%) | Cities of Helsinki, Bonn, Stockholm |
| Arrows with traffic lights | 6 (26%) | Cities of Turku, Skiathos, Barcelona |

Lastly, in terms of enhancement and obfuscation, there were two instances within two VLRs that had enhancement characteristics, and one instance in one VLR with obfuscation characteristics. For enhancement, the Cities of Hanover and Ghent had two graphs where the axis did not start from zero, thereby emphasizing the results. For obfuscation, it can be argued that the color choice decision by the City of Kiel for one of its pie charts makes it difficult to see the number and percentage of female refugees turned away as if to hide the fact. The three visualizations are seen in Figure 15 and Figure 16.

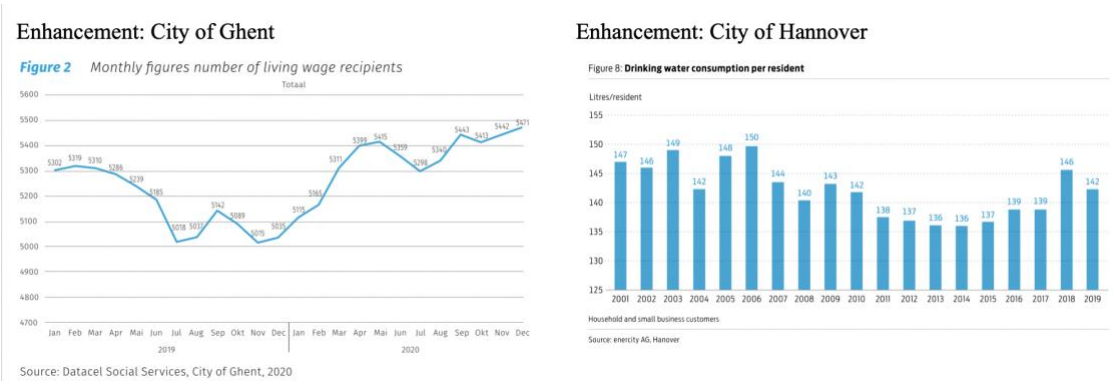



**Figure 15 Example of enhancement in the VLRs (City of Ghent, 2021, p. 18; City of Hanover, 2020, p. 43)**

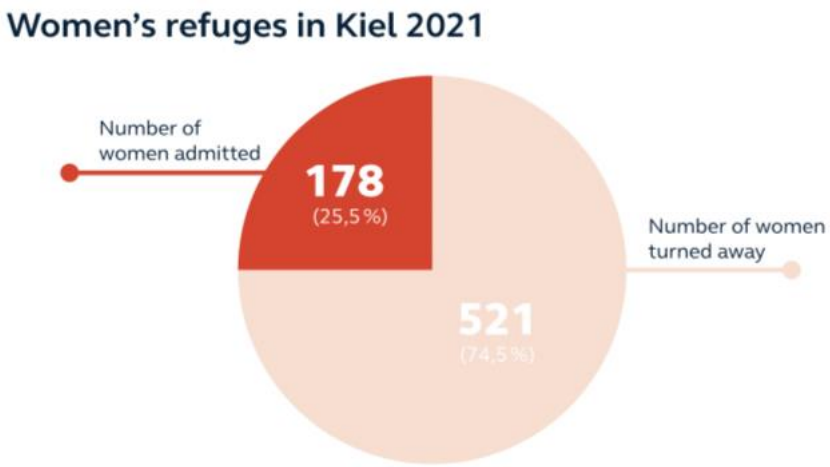



**Figure 16 Example of obfuscation in the VLRs (Landeshauptstadt Kiel, 2022, p. 30)**

### 6.2.6 Quality assurance

Quality assurance as the underlying principle that supports the process of performance measurement was observed through three main categories: general quality assurance, the identification of assurance bodies, and the legitimacy of measurement systems.

*General quality assurance*

Any mention of enacted qualities throughout the process was included for this section, and this resulted in mentions of validity, functionality, and repeatability. For example, as for validity, the City of Hanover explicitly stated, "To make sustainable development measurable, this report uses so-called sustainability indicators that must meet the following test criteria. Validity: Does the indicator appropriately reflect the content of the goal or area to be reviewed?"(City of Hanover, 2020, p. 19). In terms of functionality, the City of Turku mentioned:

> *"The City Strategy will be updated in the near future and it is then advisable to take a moment to examine the United Nations Sustainable Development Goals and set more precise goals in the city context. The City's management system will also be changed, and the City will transition from the current employment relationship-based leadership to a politically appointed mayor. This also gives an opportunity to prepare the model for the management of sustainable development in such a way that its functionality and policy consistency can be ensured in the future."* (City of Turku, 2020, p. 80)

Repeatability was found to be one of the concepts of quality assurance in previous research, and in the context of the VLRs, it can be argued that the explicit explanation of the creation process of this report can lead to increased chances of appropriate repeatability. It was previously found that 14 reports had mentioned methodology but within those 14, six reports (the Cities of Amsterdam, Espoo, Turku, Bonn, Düsseldorf, and Hanover) had step-by-step instructions on how they conducted their research. For example, the City of Espoo carried out a survey within its administration and they listed the eight questions within that survey in the VLR, and the City of Amsterdam dedicated three of its 71 pages to explaining its six-phase VLR construction process (City of Amsterdam, 2022; City of Espoo, 2020, p. 15).

*Identification of assurance bodies and legitimacy of measurement system*

A total of 14 VLRs (Cities of Düsseldorf, Malmö, Vantaa, Espoo, Turku, Bonn, Amsterdam, Hanover, Kiel, Helsinki, Helsingborg, Skiathos, the municipality of

Gladsaxe, the Generaliat Valencia) mentioned some sort of assurance bodies, whether it was for the VLR report creation, the SDG implementation process within the LRG, a general government structure, or for the implementation of specific SDG goals and targets. A summary of these classifications as well as examples taken directly from the VLRs can be seen in Table 19.

**Table 19 Summary of identified assurance bodies**

| Category | Number of VLRs (%) | VLR | Example |
|---|---|---|---|
| Within the VLR process | 9 (39%) | Cities of Gladsaxe, Helsinki, Epsoo, Turku, Amsterdam, Malmö, Düsseldorf, Bonn, Vantaa | "From this follows the decision from the alder- woman of Sustainability and the alderman of Economy to provide an administrative assignment to conduct a VLR of the city of Amsterdam" (City of Amsterdam, 2022, p. 126). |
| Within the SDG implementation process | 3 (13%) | Cities of Malmö, Helsingborg & Region of Lombardy | "A special Office for Sustainable Development was established in 2017 in order to accelerate work on the 2030 Agenda in Malmö, and this was tasked with heading, coordinating, developing, supporting and communicating the city's strategic work on the 2030 Agenda" (City of Malmö, 2021, p. 17) |
| General governmental structure | 1 (4%) | State of North-Rhine Westphalia | "The State Cabinet decides on all matters regarding the implementation and further development of the NRW Sustainability Strategy based on a cabinet draft submitted by the Minister for Climate Protection, Environment, Agriculture, Nature and Consumer Protection" (The State Government of NRW, 2016, p. 29) |
| For specific SDGs | 1 (4%) | City of Kiel | For SDG5, it stated, "In Kiel, the central contact point for all questions concerning gender equality is the Department for Gender Equality and its equality officers" (Landeshauptstadt Kiel, 2022, p. 27) |

# 7 Discussion

This section will further discuss the results that were gathered based on the conceptual framework in an attempt to answer the research question: "What is the nature of the indicators communicated within the Voluntary Local Review reports by the local and regional governments of the European Union?" guided by the following sub-questions.

1. What are the characteristics of the indicators on the EU VLRs?
2. What are the characteristics of the VLRs as public reports?

Before diving into the specific discussion of these questions, the general points will be discussed. One of the most surface-level differences between the VLRs was the number of pages and the indicators, with some reports having fewer pages but more indicators and vice versa. It should be noted that three VLRs had no mention of indicators, although Generalitat Valencia's lack of indicators could be attributed to the fact that this VLR was published before the official VLR structure was made public. First, the fact that the majority of the VLRs discussed its indicators is reflective of past research that has stated their importance in the context of sustainability and the SDGs (Lyytimäki et al., 2013; Mair et al., 2018; Ramos & Caeiro, 2010; Scerri & James, 2010). On the contrary, the lack of mention from the two VLRs could reflect the general level of knowledge about indicators, including if they are differentiated between mere data points which is what especially the City of Kiel's VLR appeared to interpret them as, and if there are understood to be information on performance measurements. Although it is probably safe to assume that the concept of indicators should be no stranger to all of these LRGs due to the history of indicators and their presence in public administration, a lack of knowledge may have been a factor towards these two VLRS as well as the lack of imposed standardization of these reports in which the LRGs may have interpreted the explicit mention of indicators to be unnecessary.

Another surface-level difference that was observed was how the LRGs introduced the SDGs in the VLRs; 12 (about 50%) of the VLRs had written about each SDGs, while 10 (about 43%) created their own dimensions that followed local strategies. Furthermore, two (about 8%) of the VLRs introduced only the HLPF-designated SDGs. Despite this disparity, all of the VLRs that used either the SDGs or local dimensions mapped out the SDGs to their local indicators or initiatives. 13 (about 56%) linked the global goals to local initiatives while eight (about 34%) linked the global targets to local initiatives. From the gathered results, it can be said that the LRGs are able to present connections between local initiatives and indicators to the global SDG goals and/or targets. There is tremendous potential for future research to analyze a correlation between these connections to effective policymaking.

Lastly, regardless of the given VLR structure by the UNDESA as well as international guidelines on report creation, the indicators, the content, and the presentation of the Voluntary Local Reviews of the EU LRGs were truly varied, even within the same country. The lack of comparability has been pointed out in previous research (Schmidt-Traub et al., 2017), and the variety seen through the results of this research seemed to echo this point. In general, the variability can be a reflection of the differences within the local and regional governments, which motivated the creation of the VLR format to highlight these differences in the first place (United Nations, n.a.-f), making the variability warranted and appropriate. Therefore, the following sections will highlight these differences for hopefully practical application to further the VLRs in more LRGs in the EU.

In this section, a discussion about each of the six steps of the performance measurement process will be conducted.

## 7.1 What to measure: the implications of predominant learning objectives, a universal public value, and agency stewardship

The first step within the performance measurement process was the identification of priorities and context as deduced by objectives, public value, and influences of agency stewardship as per the framework. In the end, the objective of the VLR could be divided into four categories: to give account, to steer and control, to learn, and to communicate. Learning was the most popular objective, follow by to steer and control next, then to give account, and finally to communicate. The implications of the prominence of certain objectives relate back to how the usage of performance information is tied to the coupling of measurement to judgment, and the impact it may have on the organization as seen in Figure 3 (Van Dooren et al., 2015, p. 123). The objective of learning is the most loosely coupled to judgment, meaning that the results do not completely reflect the actions that led to the measurement. Having a learning objective means allowing room for interpretation of the results as well as a discussion mainly within the organization that consequently leads to having a lower impact on the organization (Van Dooren et al., 2015, p. 123). The results of this research with the prominence of “learning” objectives may reflect the acknowledgment of inclusivity required to create this report, as it spans multiple teams within each LRG. In the day-to-day operations, these teams may operate in relative silos but the VLRs serve as an opportunity to synthesize these different aspects for the first time and learn, as it was stated in some of these VLRs.

The prominence of the learning objective and its implications further reflect the nature of the VLRs and the SDGs as voluntary; although there may be national and local strategies

that incorporate the SDGs into policies, the lack of stress on creating these VLRs for accountability reasons may be signaling that these VLRs are not a report on the past performance of SDG implementation as much as they are about the learnings that can be gained for the future by creating this report. After all, it is said that "in performance measurement, 'voluntary' means it will not be done" (Charbonneau, 2011). Thinking back to the official definition of the VLRs as a reporting tool and a process for political and social commitment (United Cities and Local Governments & UN-Habitat, 2021), the results of this research show these LRGs are using this opportunity to learn and orient themselves towards the future rather than controlling the present or analyzing the past. However, it may also be said that orienting performance information for accountability usage and analyzing the past to understand where performance fails can be crucial in order to properly orient their future (Van de Walle & Van Dooren, 2009). On the other hand, prior research of accountability-oriented policies and programs has concluded the specific goals, plans, and results caused by this focus have led to "administrative constraint" and "elevated conflict among multiple levels of program management", along with measuring for the sake of measuring as indicators become more disconnected to actual goals (Heinrich, 2002). Therefore, it is key to have an underlying source of trust between members of public administration and government.

In terms of public value, all of the reports mentioned underlying values that fueled the creation of the report, which is reassuring to know that the LRGs were creating this for the public rather than for personal reasons like international recognition. The universal public value - "leave no one behind" – has penetrated about a third of the LRGs which may be interpreted to be good or bad; if this value is considered "universal", it would be assumed that ideally all of the VLRs should mention this. On the other hand, these cities have their own values that they prioritize, which once again reflects the nature of localizing the SDGs. The significance of public value in indicators and performance measurements was that it aligns the public value to the aims and objectives of the organization conducting this VLR.

Finally, the hints of agency and stewardship relationships were noticed as well. Reflecting on the literature review that found agency relationships are more top-down with hard goals that establish incentives while steward relationships are based on bottom-up goal creation with no incentives as the relationship is based on trust, it was difficult to decipher from the VLRs alone precisely what kind of a relationship these LRGs have with their higher levels of government. However, it was assumed that some of these LRGs have an agency relationship because there were sub-national or national-level projects that enforced these reports. Additionally, some of the VLRs pointed to the out-of-scope nature of some of the indicators for the LRGs, which could also hint at the indicator being

nationally enforced but with no available data for the LRG. From the VLRs alone, the specifics of the relationship between the LRG and their agency is unconclusive, but future research could look into the threat of moral hazard or measure fixation if the LRGs were to have an agency relationship.

## 7.2 Indicator selection: descriptive indicators for learning purposes

The conceptual framework laid out four parts to an indicator: dimension, complexity, position, and stage along with their sub-categories. It resulted in the majority of the indicators being time series, descriptive, leading when there was context, and output/outcome based.

First, having a time series-based indicator reflects the uniformity over time which is one quality that makes for a "good" indicator (Bouckaert, 1993). In addition, having a time series-based indicator could also be used to assess the fit for policy initiatives as the fixed variables with time changes allow for comparison for such (Van Dooren et al., 2015). On the other hand, the fewer amount of cross-section and cross-section-and-time-series indicators may be reflective of the lack of detailed data collection that prohibits a cross-section presentation. A cross-section indicator would allow for more insight into different attributes, most commonly age, gender, or other cities. Additionally, comparisons within and across sectors would allow for a comparison of results of policies and management (Van Dooren et al., 2015). Especially considering the universal public value of "leave no one behind" and the emphasis on representation, the presence of cross-section indicators could avoid the potential under-representation that often is an issue when considering indicators albeit at the expense of data collection practicalities (Gudmundsson et al., 2016). At the same time, the consideration of more data could lead to an information overload that could "reduce the efficacy of political institutions" (Congleton & Sweetser, 1992).

Second, the dominance of descriptive indicators over ratio, normative and composite indicators show the straight-forwardness of these indicators as measurable indicators. However, the lack of ratio indicators that otherwise could assess relative improvements or compare different-sized entities, and the lack of normative indicators that show the gap between a goal and the status quo can mean that fewer questions can be answered with these descriptive indicators about the current situation. Indicators are said to be applied to eight applications: to describe, to forecast, to review, to diagnose, to decide, to account, to learn, and to communicate (Gudmundsson et al., 2016). Descriptive indicators can be applied to perhaps two of those applications, namely, to describe and to review. The expansion in the type of indicators used could also expand its applications. However, considering the results and discussion of the VLRs as mainly a tool for "learning" rather

than accountability that could perhaps be expressed via normative indicators through assessing progress, or to steer and control that could be expressed via ratio indicators through comparisons of different variables to understand in which direction requires action, it makes logical sense for the majority of indicators used in the VLRs to be descriptive. Furthermore, the lack of composite indicators means that these VLRs are less at risk for the pitfalls of composite indicators, such as over-generalizations. In summary, the emphasis on the aforementioned time series indicators and the descriptive indicators here reflected similar findings to past research that analyzed Québec City's performance measurement framework that found municipal managers "to not compare themselves to others, not to set targets for themselves" as well as even not using the indicators, and not including them in the reports (Charbonneau, 2011). This study further found the results of performance measurements to be guiding the objective for performance over the span of six years, rather than the other way around. As the VLRs are still relatively young, this research by Charbonneau may provide hints for what may be coming in the future.

Third, the leading or lagging aspect of indicators can only be assessed with context, and for VLRs that contained enough context, most were leading indicators. The majority of these VLRs did not provide sufficient context for this researcher to assess these two classifications. (This could have been avoided by changing the research design to include complementary interviews, as past research that analyzed indicators in an SDG context had done (Lyytimäki et al., 2020).) Simultaneously, this research's classification of leading and lagging was made with the available context and should be taken with caution, especially as these cause-and-effect categorizations can be difficult to understand from an outside perspective. That being said, the significance of leading and lagging indicators is of course their usage to forecast and analyze the past, but more importantly, it may signal a clear reason for having that indicator. For example, for the indicator "Population-weighted linear distance to the nearest primary school", the City of Bonn reasoned for the importance of this indicator by explaining, "The closure of elementary schools can lead to a reduction in the use of public transport services, which in turn favors private motorized transport" (City of Bonn, 2022). Compared to other reports where indicators would be listed without much context, reports like the City of Bonn's were more convincing even if it did not specify if these cause-and-effects happen in reality.

Finally, the majority of the indicators were output or outcome indicators, and not input. This follows the general movement within indicator usage to focus on outputs and outcomes instead of managing inputs, as it had been done so in traditional government bureaucracy (Verhoest, 2017). Furthermore, individual VLRs had mentioned the restrictions of measuring input indicators, "Purpose of the indicator: What does it track? While "outcome indicators" and "impact indicators" are good at gauging relevance to

sustainability, the informative value of "output indicators" and especially "input indicators" is limited. "Powerful" indicators in particular are often not (yet) available" (City of Hanover, 2020, p. 19).

### 7.3 Data collection: the possibility for reaping the benefits of digitalization

The degree of data collection was analyzed based on the data (un)availability, availability and location of data sources, and the aggregation/disaggregation and sampling of data and data sources.

First, in terms of data unavailability, two main categorizations were identified in the results: not having the data at all or having the data for some years but not for others. The former may signal the city's acknowledgment of the importance of these qualities, and the unavailability of data should not hinder the creation of the indicator, but a "good" indicator is also measurable and feasible to measure thus a plan for future collection may have been better as it was not explicitly stated in the VLRs. Additionally, the results show that changes within the city administration system had impacted data collection, and this reflects the dependence of these indicators on greater city strategies and/or data infrastructure. As these local systems will inevitably change in the future, as well as potentially sub-regional and national systems, this may also have future implications on data collection that in turn will have an impact on setting the indicators themselves. The influences of digitalization like these on the SDGs in general is still a young field but considering the widespread effects of digitalization (Castro et al., 2021), it may be imperative for future research to further observe these interactions. However, it is noted that none of the other VLRs mentioned the impact of system changes or any effects of digitalization on data collection.

Second, in terms of the data sources and their indication, the current results show that about half of the VLRs do not state data sources but for the ones that do, a wide variety of data sources are used from external international and national databases to internal data sources. The variety in the levels of different data sources has its advantages and disadvantages, as reflected in Van Dooren et al.'s research on performance management; internal sources tend to be cheaper and more readily available than external ones (Van Dooren et al., 2015, p. 71). In addition, there were certain indicators that would have multiple sources both internal and external, pointing to the possibility of triangulating the data that could potentially be a reflection of a principal-agent relationship where the principal may not trust the agent's data (Van Dooren et al., 2015, p. 71). Moreover, a trend seen in the LRGs of the same country was the use of centralized data sources, mainly in Finland and Sweden.

The aforementioned digitalization of data availability also affects data sources; in the Swedish cities of Helsingborg and Malmö, they provided the ID numbers from their official statistical institution called Kolada on the VLR. Upon further examination of this service, it was found that Kolada is an open and free database supported by municipal and organizational unit data, and this platform is half-owned by the Swedish State that presents and visualizes various key indicators that were created not specifically for the SDGs but in general to measure the "quality and performance of Swedish county councils and municipalities" (RKA, n.d.; Vikström, n.d.). Malmö's VLR mentioned, "The City of Malmö has access to extensive datasets collected as long time series, which provides good opportunities to define an optional baseline" about Kolada, reflecting the benefits of using such a platform (City of Malmö, 2021, p. 33). These kinds of nationwide data platforms were not mentioned in the other VLRs, albeit the mention of national data sources was seen in the Finnish cities, which shows the future potential for other countries to join forces with their municipalities for data collection and provision to not just facilitate report creation for the LRGs but also to unify understanding of these indicators as well as conducting transparent communication to citizens. This is echoed by research conducted on the impacts of digitalization that stated that a more digitalized environment for data collection could lead to wider availability of data sources and more frequent data collection, as well as the potential to have an option of aggregating or disaggregating data when necessary (Castro et al., 2021). Conversely, having these centralized databases with pre-given indicators may unintentionally put the LRGs at risk in their ability to create their own locally relevant indicators and collect locally relevant data (Van de Walle & Van Dooren, 2009). However, it was seen that the Swedish LRGs use Kolada and utilize various data sources for local indicators, allowing them to not be entirely reliant on Kolada as well as vary their data sources, thus evading the aforementioned risk.

Furthermore, disaggregation and sampling of data and data were rarely mentioned. It is difficult to assess the reasons behind this; it may be that this did not take place, or that the details did not make it on the VLR. Future research could approach both the indicator creators and the report creators to see where there may be discrepancies that ultimately led this to not be written in the VLRs.

### 7.4 Analysis: transparency of report creation process

The analysis part of the performance measurement process was analyzed based on the availability of methodology, which was highlighted due to the potential reflection of the researcher's biases, views, and assumptions. In the end, the methodology varied greatly in terms of the existence of the methodology as well as how it was explained. Although the majority of the VLRs had a 'methodology' section, the degree ranged from having a

detailed step-by-step process of how the reports and indicators were selected, to bullet points summarizing guidelines and sources it used. Ultimately, these potential biases, views, and assumptions could not be concluded from the stated methodology alone.

## 7.5 Report: the lack of guidelines but no unnecessary embellishments or masking

Despite the proposed VLR structure by the UNDESA, there was a variety of lengths and thoroughness to the VLRs. However, it can also be said the VLRs written before the proposed structure lacked key components that make a VLR, namely the objective of the report as well as the methodology. This attests to the importance of having a pre-defined structure and the potential for further brush-up of the structure, although there was a clear lack of use of international guidelines such as the IIRF and the GRI as reflective of past research. An initial argument for the lack of international guideline usage may have the distinction between sustainability and sustainable development; the IIRF and the GRI refer to sustainability reporting and not "sustainable development" reporting. However, this argument can be refuted by examining how the VLRs present these two concepts. For example, the City of Amsterdam's VLR that mentioned the commencement of the VLR creation process was given by the "Alderpersons of Sustainability and Economy", and Bonn stated, "With this VLR, we are testing a new format for our sustainability reporting, together with ten other municipalities in North Rhine-Westphalia" (City of Amsterdam, 2022, p. 31; City of Bonn, 2022, p. 5). Therefore, it can be said that these LRGs regard the concepts of sustainability and sustainable development as synonymous, thus the principles and guidelines that these international organizations speak of could apply to these "sustainability" reports as well. Past research has also shown the lack of adoption by public institutions of these guidelines, and so the findings of this research further press the need for the coordination between these guideline-setting organizations who are the experts on 'good' reporting, the United Nations that spearhead the SDGs, the countries and its local and regional governments that ultimately implement the SDGs and have the responsibility for reporting on them.

In terms of visualizations, 17 of the reports used some sort of visualization while the rest did not utilize any visualizations. These remaining six reports reflect the lack of usage supported by previous research that proved public reports to be textually complicated without any visual aids in an attempt to explain the details of sustainability. On the other side, when there were visualizations, it was neither exaggerated nor intentionally hiding unfavorable results that were seen in previous research of private sector reports. The prevalence of these impression management techniques in the private sector is understandable since legitimacy theory in this context would imply these private sector

firms having good relationships with their stakeholders while following regulations but fully utilizing whatever they can. Inversely, although the same legitimacy theory applies for public institutions, it may be said that these results show this theory does not apply to the degree that it does to private sector firms therefore public institutions do not have to rely on impression management methods as much. This calls for further research into the use of impression management through visualizations in public reports, which to the researcher's knowledge is still lacking.

## 7.6 Quality assurance: the existence of a foundation

Quality assurance was analyzed based on general qualities that were mentioned in the VLRs as underlying values of the report creation process. These included validity, functionality, and repeatability as guided by past research. Additionally, the notation of related stakeholders who assessed the performance measurement process was reflective of past research. To have observed these qualities in the VLRs could imply that these LRGs understand what is necessary for performance measurements, and that the foundation (regardless of its implementation, execution, or results) for 'good' performance management is there for those LRGs. On the other hand, these qualities were not present in all of the VLRs which suggests room for knowledge sharing to create a common understanding. By having a based foundation for the performance measurement process, the qualities of the statistical and the production process of indicators can be checked (Van Dooren et al., 2015). Without one, it could lead to non-use, misuse, or over-use of performance information, and inappropriately set indicators that are fed by inaccurate data sources thus leading to wrong analysis of the status quo which all ultimately can lead to an erosion of trust from citizens.

# 8 Conclusion

This research intended to fill the gap in understanding the nature of the local SDG global indicators within the EU LRGs by examining the local indicators through the lens of the performance measurement process. In doing so, this research hoped to contribute to how these LRGs are managing performance, and ultimately this research was the first step in analyzing this from a process perspective. In the following sections, the answers to the aforementioned research and sub-questions will be discussed along with the implications, limitations, and significance of this thesis. Finally, potential future avenues of research will be explored.

## 8.1 Answer to Research Question

The following main research question laid the foundation for this research: "In what way are performance indicators for the sustainable development goals presented in the Voluntary Local Reviews by the cities of the European Union?" Through the literature review, a conceptual framework devising a frame for analyzing indicators was created using the performance measurement process. The qualitative case studies of 23 Voluntary Local Reviews by the local and regional governments of the European Union were guided by two additional sub-questions stemming from the main research question: 1. What are the characteristics of the indicators on the EU VLRs? 2. What are the characteristics of the VLRs as public reports? The results of these questions uncovered current performance information usage and insight into what could be done in the future to guide SDG achievement by the LRGs.

Ultimately, there was no straightforward answer to all of these questions as it depended on each case. However, the indicators were largely based on descriptions rather than measuring against norms, and these indicators were provided with relatively little context to accurately assess if these indicators were provided to determine the result of intended actions or to determine forecasting. Additionally, these descriptive indicators were monitored over time rather than through comparisons with other variables or cities and countries. Furthermore, the process of collecting the data to produce the results of these indicators was not as complicated as previous research had uncovered with potential pitfalls of aggregating and/or disaggregating and sampling data and data sources, as these VLRs did not mention that much about these problems. However, the lack of details on these points could have been a problem with the report itself, as the nonmention of these details may have just been due to the lack of unavailable space or other unknown reasons. On the other hand, it was found that certain countries had a central database that allowed some of these VLRs to use the data and indicators from that database for their reports,

reflecting the potential for other LRGs and nations to follow suit and utilize more digital technologies for not just the VLR creation but for indicator creation. As for the analysis and methodologies of the indicators and the VLRs, the general lack of details in all of the VLRs made it difficult to accurately assess the state of methodologies in the indicator and VLR creation process, thus opening up avenues for future research that could include interviews with relevant stakeholders. Finally, analyzing the VLRs as public reports disproved past research that had analyzed private sector firms' sustainability reports that manipulated their reports to fill their legitimacy gap through methods such as enhancement and obfuscation. Fortunately, these VLRs did not exhibit common methods for impression management, but unfortunately, the lack of mention of international guidelines for sustainability reports such as the GRI and IIRF echoed previous research on public institutions' insufficient and absence of following these guidelines.

The aim of the SDG global indicators is, "help countries plan their policies, measure their progress, find gaps and report the results to the international community" (SDSN Youth, n.d.). Furthermore, the VLR's official aim by the United Nations is to "both a reporting tool to assess, monitor and present local achievements in implementing the SDGs, as well as a process to enhance and expand the political and social commitment of a variety of stakeholders to the SDGs, orienting local priorities and development planning" (United Cities and Local Governments & UN-Habitat, 2021). These two aims both have in common the characteristic of planning policies, assessing progress, and reporting the results. On the other hand, the aims, or uses, of performance measurement as defined in past theories are three-fold: to learn, to steer and control, and to give account. It can be said that the aims of the global indicators and the VLR overlap with uses of performance measurement; assessing progress can be synonymous with learning for the future, creating and planning policies can be synonymous with steering and controlling the present, and reporting is synonymous with taking account by communicating the performance of the past. In this way, the results from this research show that the current state of indicators and reporting as seen in the VLRs is mostly based on learning for the future.

## 8.2 Implication of Findings

The implications of the results of this thesis can be applied to the field of local SDG indicators, and to the field of research on performance information usage. First, in terms of the former, this research may help scholars, policymakers, and citizens alike that the current ways that LRGs create their reports and present their indicators point to a learning objective that usually is associated with soft uses that have lower impacts on an organization. As explained in Section 7.1, soft use allows for discussions to interpret

results and consequent actions as opposed to hard uses where a 1:1 relation between the measurements and the results and actions can be observed, and this kind of hard use often leads to stricter judgment where the results and actions have a higher impact on the organization. On the other hand, soft use accommodates the inclusiveness needed to understand the problems and solutions to meet the SDGs as the SDGs are a wicked problem that makes it difficult to pinpoint details, but it is also known that using evidence-based management in general for wicked problems is difficult precisely for the aforementioned reason (Van de Walle & Van Dooren, 2009). In addition, soft use also leads to having a lower impact on organizations with lesser judgment. All of this is in line with the voluntariness of the SDGs, their indicators, and the VLRs. However, considering the stressed importance of the SDGs and their indicators on all levels of society in order to achieve the 2030 goal, there may be more demand for these indicators and reports to lean toward hard uses that will make these public institutions susceptible to harsher judgment but also lead them to steer and control, or to show accountability for their actions.

Second, in terms of academic and theoretical contributions, this research was conducted using the pre-established performance measurement process as a conceptual framework and adding other theories together, thus aimed to contribute to advancing existing theories through results that have mirrored past researched that utilized the same theories. However, there were certain aspects of this framework that were found to be inconclusive due to the nature of the methodology and data sources, thus future research using this framework should consider multiple data sources and a combination of different types of methods such as content analysis and interviews to attain conclusive results.

## 8.3 Significance of Thesis

To the extent of the researcher's knowledge, this was the first attempt at analyzing the indicators of local SDG indicators of the LRGs in the EU as a part of the performance measurement process. Past research has also analyzed indicators of the EU LRGs, but it utilized the SDGs as an analytical framework and their aim was to provide a follow-up to the European Handbook on SDG Voluntary Local Reviews (Ciambra et al., 2021). The uniqueness of this thesis lies in its focus on performance management, the performance measurement process, and the focus on performance measurement usage. In doing so, it attempted to contribute to how these LRGs are using performance measurements and information towards the SDGs.

This research can contribute to the ever-growing topic of Sustainable Development Goals and their reporting, especially as new kinds of reporting such as the EU-level Voluntary Review of the SDG implementation being presented at the 2023 UN High-level Political

Forum on Sustainable Development in July 2023 (European Commission, 2022). This can aide future creation as well as past reflections of the local indicators of SDGs that can ultimately lead to more resilient communities while empowering local institutions (ElMassah & Mohieldin, 2019).

## 8.4 Limitations of Thesis

Although the objectives were met and the research question was answered, there were several limitations to this research that should be considered in combination with the results that allow for future potential research. Due to the time span of this research, the data source was limited to one type of document, the VLRs, and this may have inhibited perspective on the situation due to possible incomplete information. This research attempted to cover possible biases that the LRGs may have transposed onto the VLRs through an examination of VLR as reports, but there are further risks of bias. Furthermore, the lack of time series comparisons of the VLRs that only describes a situation at a certain moment may have limited the findings that otherwise may have found further results.

## 8.5 Potential Future Research Avenues

There are several potential future research avenues that have been identified through this research. When examining pre-existing research on VLR analysis, the scope of the summaries was intentionally limited in order to align itself with the scope of this research i.e. the indicators. However, indicators are only one tool to aid in the localization of the SDGs, and various other components such as the institutionalization of the SDGs and the VLRs, involvement of civil society, and others are equally as important to ensure the operationalization of sustainability into society (Hezri & Dovers, 2009). Therefore, future research can examine the relationship between the different variables and understand the correlation between influencing factors of successful SDG localization.

Additionally, this research did not look into how the indicators as performance information were used for policymaking as this could not be deciphered by the VLRs alone. However, as previously mentioned, performance information can only be deemed a success if it is actually being used. Therefore, future research could focus on the usage of these indicators in policymaking or initiatives. This is worth considering especially since it is known that the PI may not contribute to reducing uncertainty, may not automatically translate into usage as the providers of PI and the potential users may be different, and standardized indicators may inversely restrict its usefulness to an organization (Van de Walle & Van Dooren, 2009).

The basis for having these VLRs is to enhance comparability, which is essential to further the SDGs, but this research has concluded that the indicators and the reporting are varied. Therefore, future research could explore potential frameworks that are able to frame the VLRs in a more comparative way. However, as the societal problems that the SDGs cover, ranging from climate change to poverty to healthcare, become ever more dire, it is expected that further research into this topic will be conducted. Therefore, this research hoped to have contributed as a step into the academic research of localization of performance measurements of these wicked problems.

## References

Adams, C. A., & Evans, R. (2004). Accountability, Completeness, Credibility and the Audit Expectations Gap. *The Journal of Corporate Citizenship*, *2004*(14), 97–115. https://doi.org/10.9774/GLEAF.4700.2004.su.00010

Afolabi, H., Ram, R., & Rimmel, G. (2022). Harmonization of Sustainability Reporting Regulation: Analysis of a Contested Arena. *Sustainability*, *14*(9). https://doi.org/10.3390/su14095517

Allen, C., Reid, M., Thwaites, J., Glover, R., & Kestin, T. (2019). Assessing national progress and priorities for the Sustainable Development Goals (SDGs): Experience from Australia. *Sustainability Science*, *15*, 521–538.

Ballard, A. (2020). Promoting Performance Information Use Through Data Visualization: Evidence from an Experiment. *Public Performance & Management Review*, *43*(1), 109–128. https://doi.org/10.1080/15309576.2019.1592763

Bandola-Gill, J., & Smith, K. E. (2022). Governing by narratives: REF impact case studies and restrictive storytelling in performance measurement. *Studies in Higher Education*, *47*(9), 1857–1871. https://doi.org/10.1080/03075079.2021.1978965

Basque Government. (2021). *5th Monitoring Report: 2030 Agenda Basque Country 2021*.

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

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

Bird, S. M., Cox, D., Farewell, V. T., Goldstein, H., Holt, T., & Smith, P. C. (2005). Performance Indicators: Good, Bad, and Ugly. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *168*(1), 1–27.

Bjurstrøm, K. H. (2020). Principal–Agent or Principal–Steward: How Ministry–Agency Relations Condition the Impact of Performance Management in the Steering of Government Agencies. *Public Performance & Management Review*, *43*(5), 1053–1077. https://doi.org/10.1080/15309576.2020.1739087

Boiral, O., Heras-Saizarbitoria, I., & Brotherton, M.-C. (2019). Assessing and Improving the Quality of Sustainability Reports: The Auditors' Perspective. *Journal of Business Ethics*, *155*(3), 703–721. JSTOR.

Bora, I., Duan, H. K., Vasarhelyi, M. A., Zhang, C. (Abigail), & Dai, J. (2021). The Transformation of Government Accountability and Reporting. *Journal of*

*Emerging Technologies in Accounting*, *18*(2), 1–21. https://doi.org/10.2308/jeta-10780

Bouckaert, G. (1993). Measurement and Meaningful management. *Public Productivity and Management Review*, *17*(1). https://doi.org/10.2307/3381047

Boyne, G. A., Farrell, C., Law, J., Powell, M., & Walker, R. M. (2003). *Evaluating Public Management Reforms*. Open University Press.

Bozeman, B. (2002). Public-Value Failure: When Efficient Markets May Not Do. *Public Administration Review*, *62*(2), 145–161.

Bozeman, B. (2007). *Public Values and Public Interest: Counterbalancing Economic Individualism.* http://press.georgetown.edu/book/georgetown/public-values-and-public-interest

Brugmann, J. (1997). Is there a method in our measurement? The use of indicators in local sustainable development planning. *Local Environment*, *2*(1), 59–72. https://doi.org/10.1080/13549839708725512

Budding, T., Faber, B., & Schoute, M. (2022). Integrating non-financial performance indicators in budget documents: The continuing search of Dutch municipalities. *Journal of Public Budgeting, Accounting & Financial Management*, *34*(1), 52–66. ProQuest Central. https://doi.org/10.1108/JPBAFM-02-2020-0009

C. Edgar, V., Beck, M., & Brennan, N. M. (2018). Impression management in annual report narratives: The case of the UK private finance initiative. *Accounting, Auditing & Accountability Journal*, *31*(6), 1566–1592. https://doi.org/10.1108/AAAJ-10-2016-2733

Castro, G. D. R., Fernández, M. C. G., & Colsa, Á. U. (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

Cavalcante, R., Rodrigues, R., & Lagioia, U. (2019). Public Sector (Un)Sustainability: A study of GRI adherence and sustainability reporting disclosure standards in Public Institutions and State-Owned Companies of the Public Agency Sector. *Cuadernos de Contabilidad*, *20*, 1–28. https://doi.org/10.11144/Javeriana.cc20-49.psss

Charbonneau, É. (2011). Assessing the effects of an Intelligence Performance Regime: Quebec's Municipal Management Indicators, 1999–2010. *International Review of Administrative Sciences*, *77*(4), 733–755. https://doi.org/10.1177/0020852311419390

Cho, C. H., Michelon, G., & Patten, D. M. (2012b). Enhancement and obfuscation through the use of graphs in sustainability reports: An international comparison. *Sustainability Accounting, Management and Policy Journal (Print)*, *3*(1), 74–88. https://doi.org/10.1108/20408021211223561

Cho, C. H., Michelon, G., & Patten, D. M. (2012a). Impression management in sustainability reports: An empirical investigation of the use of graphs. *Accounting and the Public Interest*, *12*(1), 16–37. https://doi.org/10.2308/apin-10249

Choi, I., & Gil-Garcia, J. R. (2022). Do different presentations of performance information on government websites affect citizens' decision making? A survey experiment. *International Public Management Journal*, *25*(1), 140–158. https://doi.org/10.1080/10967494.2021.1913459

Ciambra, A., Siragusa, A., & Proietti, P. (2021). *European SDG Voluntary Local Reviews: A comparative analysis of local indicators and data* (Scientific Analysis or Review KJ-05-21-093-EN-N (online)). Publications Office of the European Union. https://doi.org/10.2760/9692 (online)

Cities Alliance. (2015). *Sustainable Development Goals and Habitat III: Opportunities for a Successful New Urban Agenda* (No. 3; Cities Alliance Discussion Paper). Cities Alliance. https://www.citiesalliance.org/sites/default/files/Opportunities%20for%20the%20New%20Urban%20Agenda.pdf

City of Amsterdam. (2022). *Voluntary Local Review 2022: Impact of the Sustainable Development Goals on the City of Amsterdam*.

City of Bonn. (2022). *Voluntary Local Review 2022: The Implementation of the UN Sustainable Development Goals in the City of Bonn*.

City of Düsseldorf. (2022). *Voluntary Local Review 2022: The Implementation of the UN Sustainable Development Goals in the City of Düsseldorf*.

City of Espoo. (2020). *VOLUNTARY LOCAL REVIEW: IMPLEMENTATION OF THE UNITED NATIONS' SUSTAINABLE DEVELOPMENT GOALS 2030 IN THE CITY OF ESPOO*.

City of Ghent. (2021). *The Ghent Sustainability Report 2021: Focus on People*.

City of Hanover. (2020). *Hanover on the Path to Sustainability*. https://unhabitat.org/sites/default/files/2022/09/hanover_2020_en-optimized-optimized.pdf

City of Helsingborg. (2021). *Voluntary local review of Helsingborg*.

City of Helsinki. (2021). *From Agenda to Action: Implementation of the UN Sustainable Development Goals in Helsinki*.

City of Malmö. (2021). *Voluntary Local Review City of Malmö 2021: A review of the city's steering towards the Sustainable Development Goals*.

City of Mannheim. (2019). *THE IMPLEMENTATION OF THE UNITED NATIONS' SUSTAINABLE DEVELOPMENT GOALS IN MANNHEIM 2030*.

City of Tampere. (2022). *City of Sustainable Action: The Voluntary Local Review of the UN Sustainable Development Goals in Tampere 2022*.

City of Turku. (2020). *A Voluntary Local Review 2020: The implementation of the 2030 Agenda for Sustainable Development in the City of Turku*.

City of Vantaa. (2021). *SUSTAINABLE VANTAA BELONGS TO EVERYONE: Implementation of the UN Sustainable Development Goals in Vantaa 2021*.

Clarke, D. P., Hrasky, S. L., & Tan, C. G. T. (2009). Voluntary Narrative Disclosures by Local Governments: A Comparative Analysis of the Textual Complexity of Mayoral and Chairpersons Letters in Annual Reports. *Australian Journal of Public Administration*, *68*(2), 194–207. https://doi.org/10.1111/j.1467-8500.2009.00630.x

Congleton, R. D., & Sweetser, W. (1992). Political deadlocks and distributional information: The value of the veil. *Public Choice*, *73*(1), 1–19.

Congress of Local and Regional Authorities. (n.a.). *Sustainable Development Goals*. https://www.coe.int/en-GB/web/congress/sdg

Cunningham, G. M., & Harris, J. E. (2005). Toward a Theory of Performance Reporting to Achieve Public Sector Accountability: A Field Study. *Public Budgeting & Finance*, *25*(2), 15–42. https://doi.org/10.1111/j.0275-1100.2005.00359.x

Cüre, T., Esen, E., & Çalışkan, A. Ö. (2020). Impression Management in Graphical Representation of Economic, Social, and Environmental Issues: An Empirical Study. *Sustainability*, *12*, 379.

Dando, N., & Swift, T. (2003). Transparency and Assurance: Minding the Credibility Gap. *Journal of Business Ethics*, *44*(2/3), 195–200. https://doi.org/10.1023/A:1023351816790

de Losada, A. F. (2021). *Comparative study on SDG monitoring systems at local and regional levels*. https://gold.uclg.org/sites/default/files/eng_sdg_monitoring_systems_final_version.pdf

Dumay, J., Guthrie, J., & Farneti, F. (2010). Gri Sustainability Reporting Guidelines For Public And Third Sector Organizations. *Public Management Review*, *12*(4), 531–548. https://doi.org/10.1080/14719037.2010.496266

Eberhardt, A., & Silveira, M. S. (2018). Show Me the Data! A Systematic Mapping on Open Government Data Visualization. *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*. https://doi.org/10.1145/3209281.3209337

ECIIA & EUROSAI. (2021). *Integrated Reporting in the European Public Sector: It's Time to Act!* https://www.eurosai.org/handle404?exporturi=/export/sites/eurosai/.content/documents/news/Final-report-Integrated-reporting.pdf

ElMassah, S., & Mohieldin, M. (2019). Digital transformation and localizing the Sustainable Development Goals (SDGs). *Ecological Economics*, *169*. https://doi.org/10.1016/j.ecolecon.2019.106490

European Commission. (2022). *EU Voluntary Review on the implementation of the 2030 Agenda for Sustainable Development*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13600-EU-Voluntary-Review-on-the-implementation-of-the-2030-Agenda-for-Sustainable-Development_en

European Commission. (n.a.). *Corporate sustainability reporting*. https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en

European Union. (2022a). *Sustainable Development in the European Union: Monitoring Report on Progress Towards the SDGs in an EU Context*. https://ec.europa.eu/eurostat/documents/15234730/15242025/KS-09-22-019-EN-N.pdf/a2be16e4-b925-f109-563c-f94ae09f5436?t=1667397761499

European Union. (2022b, September 12). *Commission joins forces with local and regional government associations to boost sustainable development*. https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5395

Farneti, F., & Guthrie, J. (2009). Sustainability reporting by Australian public sector organisations: Why they report. *Accounting Forum*, *33*(2), 89–98. https://doi.org/10.1016/j.accfor.2009.04.002

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, S., & Macleod, A. (2019). *Voluntary Local Reviews: A handbook for UK cities. Building on the Bristol experience*. University of Bristol. https://unhabitat.org/sites/default/files/2021/07/uk-cities-voluntary-local-review-handbook.pdf

Francke, W. (1995). The Evolving Watchdog: The Media's Role in Government Ethics. *The Annals of the American Academy of Political and Social Science*, *537*, 109–121. JSTOR.

Fusco, F., & Ricci, P. (2019). What is the stock of the situation? A bibliometric analysis on social and environmental accounting research in public sector. *International Journal of Public Sector Management*, *32*(1), 21–41. https://doi.org/10.1108/IJPSM-05-2017-0134

Generalitat Valenciana. (2016). The Region of Valencia and the local implementation of the SDGs: A region committed to Cooperation and the 2030 Agenda for Sustainable Development.

Gladsaxe Municipality. (2022). *Voluntary Local Review from Gladsaxe 2022*.

Gobierno Autónomo Municipal de La Paz. (2018). *Objetivos de Desarrolo Sostenible Y Su Localización en el Municipio de La Paz*. http://sitservicios.lapaz.bo/sit/ods/ods_gamlp.pdf

Goldstein, H., & Spiegelhalter, D. J. (1996). League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *159*(3), 385–443.

Greiling, D., Traxler, A. A., & Stötzer, S. (n.d.). Sustainability reporting in the Austrian, German and Swiss public sector. *International Journal of Public Sector Management*, *28*(4/5), 404–428.

GRI. (2022). *GRI 1: Foundation 2021*.

Gudmundsson, H., Hall, R. P., Marsden, G., & Zietsman, J. (2016). Indicators. In H. Gudmundsson, R. P. Hall, G. Marsden, & J. Zietsman (Eds.), *Sustainable Transportation: Indicators, Frameworks, and Performance Management* (pp. 137–169). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46924-8_6

Guthrie, J., Petty, R., Yongvanich, K., & Ricceri, F. (2004). Using content analysis as a research method to inquire into intellectual capital reporting. *Journal of Intellectual Capital*, *5*(2), 282–293. https://doi.org/10.1108/14691930410533704

Hasselblad, A. (2021). What is Actually Measured? Investigating the Correspondence Among Goals and Performance Indicators in a Swedish Municipality. *2021 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 598–602.

Heinrich, C. J. (2002). Outcomes–Based Performance Management in the Public Sector: Implications for Government Accountability and Effectiveness. *Public Administration Review*, *62*(6), 712–725. https://doi.org/10.1111/1540-6210.00253

Hezri, A. A., & Dovers, S. R. (2009). Australia's Indicator-Based Sustainability Assessments and Public Policy. *Australian Journal of Public Administration*, *68*(3), 303–318. https://doi.org/10.1111/j.1467-8500.2009.00641.x

Hildon, Z., Allwood, D., & Black, N. (2012). Impact of format and content of visual display of data on comprehension, choice and preference: A systematic review. *International Journal for Quality in Health Care*, *24*(1), 55–64. https://doi.org/10.1093/intqhc/mzr072

Hood, C. (1991). A Public Management for All Seasons? *Public Administration*, *69*, 3–19.

Hood, C. (2007). Public Service Management by Numbers: Why Does it Vary? Where Has it Come From? What Are the Gaps and the Puzzles? *Public Money & Management*, *27*(2), 95–102. https://doi.org/10.1111/j.1467-9302.2007.00564.x

ICLEI. (2015). *Cities and the Sustainable Development Goals* (Urban Issues). Local Governments for Sustainability. https://www.local2030.org/library/232/ICLEI-Briefing-Sheets-02-Cities-and-the-Sustainable-Development-Goals.pdf

Inter-Agency and Expert Group on & Sustainable Development Goal Indicators. (2019). *Data Disaggregation and SDG Indicators: Policy Priorities and Current and Future Disaggregation Plans*. https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3a-Data-Disaggregation-E.pdf

IRFS. (2021). *INTERNATIONAL <IR> FRAMEWORK*. https://www.integratedreporting.org/wp-content/uploads/2021/01/InternationalIntegratedReportingFramework.pdf

ITU & DIAL. (2019). *SDG Digital Investment Framework: A Whole-of-Government Approach to Investing in Digital Technologies to Achieve the SDGs*. https://www.itu.int/pub/D-STR-DIGITAL.02-2019

J. Heinrich, C. (2007). Evidence-Based Policy and Performance Management: Challenges and Prospects in TwoParallel Movements. *The American Review of Public Administration*, *37*(3), 255–277. https://doi.org/10.1177/0275074007301957

Jones, M., Melis, A., Gaia, S., & Aresu, S. (2018). Does graphical reporting improve risk disclosure? Evidence from European banks. *Journal of Applied Accounting Research*, *19*(1), 161–180. https://doi.org/10.1108/JAAR-07-2016-0068

Kansteiner, K., & König, S. (2020). The role(s) of qualitative content analysis in mixed methods research designs. *Forum, Qualitative Social Research*, *21*(1). https://doi.org/10.17169/fqs-21.1.3412

Kaufmann, D., & Kraay, A. (2008). Governance Indicators: Where Are We, Where Should We Be Going? *The World Bank Research Observer*, *23*(1), 1–30.

Keerasuntonpong, P., Manowan, P., & Shutibhinyo, W. (2019). Reforming government public accountability: The case of Thailand. *Journal of Public Budgeting, Accounting & Financial Management*, *31*(2), 237–263. https://doi.org/10.1108/JPBAFM-05-2018-0051

Kroll, A. (2015). Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. *Public Performance & Management Review*, *38*(3), 459–486.

Kuc-Czarnecka, M., Lo Piano, S., & Saltelli, A. (2020). Quantitative Storytelling in the Making of a Composite Indicator. *Social Indicators Research*, *149*, 775–802.

Kuc-Czarnecka, M., Markowicz, I., & Sompolska-Rzechula, A. (2023). SDGs implementation, their synergies, and trade-offs in EU countries—Sensitivity analysis-based approach. *Ecological Indicators*, *146*. https://doi.org/10.1016/j.ecolind.2023.109888

Landeshauptstadt Kiel. (2022). *Voluntary Local Review Kiel 2022*.

Latapí Agudelo, M. A., Jóhannsdóttir, L., & Davídsdóttir, B. (2019). A literature review of the history and evolution of corporate social responsibility. *International Journal of Corporate Social Responsibility*, *4*(1), 1. https://doi.org/10.1186/s40991-018-0039-y

Laurent, É. (2018). Good and Bad Indicators: The Case of GDP. In *Measuring Tomorrow* (pp. 20–26). Princeton University Press. https://doi-org.kuleuven.e-bronnen.be/10.1515/9781400888634-005

Leadership Council Sustainable Development Solutions Network. (2015). *Indicators and a Monitoring Framework for the Sustainable Development Goals*. https://sdgs.un.org/publications/indicators-and-monitoring-framework-sustainable-development-goals-17958

Leavesley, A., Trundle, A., & Oke, C. (2022). Cities and the SDGs: Realities and possibilities of local engagement in global frameworks. *Ambio Volume*, *51*, 1416–1432.

Lee, M. (2002). Intersectoral Differences in Public Affairs: The Duty of Public Reporting in Public Administration. *Journal of Public Affairs*, *2*(2), 33–43.

Liguori, M., Sicilia, M., & Steccolini, I. (2012). Some Like it Non-Financial …. *Public Management Review*, *14*(7), 903–922. https://doi.org/10.1080/14719037.2011.650054

Lodhia, S., Jacobs, K., & Park, Y. J. (2012). Driving Public Sector Environmental Reporting. *Public Management Review*, *14*(5), 631–647. https://doi.org/10.1080/14719037.2011.642565

Lovre, I., Munitlak Ivanović, O., & Mitić, P. (2017). Analysis of Public Sector Efficiency in Developed Countries. *Economic Analysis (Belgrade, Serbia : 2009)*, *50*(1), 38–49.

Lyytimäki, J., Salo, H., Lepenies, R., Büttner, L., & Mustajoki, J. (2020). Risks of producing and using indicators of sustainable development goals. *Sustainable Development (Bradford, West Yorkshire, England)*, *28*(6), 1528–1538. https://doi.org/10.1002/sd.2102

Lyytimäki, J., Tapio, P., Varho, V., & Söderman, T. (2013). Use, non-use and misuse of indicators in sustainability assessment and communication. *International Journal of Sustainable Development and World Ecology*, *20*(5), 385–393. https://doi.org/10.1080/13504509.2013.834524

Mair, S., Jones, A., Ward, J., Christie, I., Druckman, A., & Lyon, F. (2018). A Critical Review of the Role of Indicators in Implementing the Sustainable Development Goals. In W. Leal Filho (Ed.), *Handbook of Sustainability Science and Research* (pp. 41–56). Springer International Publishing. https://doi.org/10.1007/978-3-319-63007-6_3

Mäkinen, K., Prutsch, A., Karali, E., Leitner, M., Völler, S., Lyytimäki, J., Pringle, P., & Vanneuville, W. (2018). Indicators for Adaptation to Climate Change at National Level—Lessons from Emerging Practice in Europe. *European*

*Environment Agency*. https://www.eionet.europa.eu/etcs/etc-cca/products/etc-cca-reports/tp_3-2018

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. https://doi.org/10.1108/JPBAFM-03-2020-0037

May, J. J., & Sanders, E. T. W. (2013). Beyond Standardized Test Scores: An Examination of Leadership and Climate as Leading Indicators of Future Success in the Transformation of Turnaround Schools. *Journal of Urban Learning, Teaching, and Research*, *9*, 42–54.

Melis, A., & Aresu, S. (2022). Analyst Following, Country's Financial Development, and the Selective Use of Graphical Information in Corporate Annual Reports. *International Journal of Business Communication*, *59*(4), 459–484. https://doi.org/10.1177/2329488419829886

Merriam-Webster. (2023). *Indicate*. https://www.merriam-webster.com/dictionary/indicates

Merry, S. E. (2019). The Sustainable Development Goals Confront the Infrastructure of Measurement. *Global Policy*, *10*(S1), 146–148.

Micheli, P., & Neely, A. (2010). Performance Measurement in the Public Sector in England: Searching for the Golden Thread. *Public Administration Review*, *70*(4), 591–600. https://doi.org/10.1111/j.1540-6210.2010.02180.x

Montesinos, V., & Brusca, I. (2019). Non-Financial Reporting in the Public Sector: Alternatives, Trends and Opportunities. *Spanish Accounting Review*, *22*(2), 122–128.

Moon, K., Blackman, D. A., Adams, V. M., Colvin, R. M., Davila, F., Evans, M. C., Januchowski-Hartley, S. R., Bennett, N. J., Dickinson, H., Sandbrook, C., Sherren, K., St. John, F. A. V., Kerkhoff, L., Wyborn, C., & Ellison, A. (2019). Expanding the role of social science in conservation through an engagement with philosophy, methodology, and methods. *Methods in Ecology and Evolution*, *10*(3), 294–302. https://doi.org/10.1111/2041-210X.13126

Moore, M. H. (1995). *Creating Public Value: Strategic Management in Government*. Harvard University Press. https://scholar.harvard.edu/markmoore/markmoore-0

Moynihan, D. P. (2008). *The Dynamics of Performance Management: Constructing Information and Reform*. Georgetown University Press. https://ebookcentral.proquest.com/lib/tuee/detail.action?docID=547813

Municipality of Skiathos. (2020). *Voluntary Local Review 2020*.

Mustea, L., Mihit, L. D., & Lobont, O.-R. (2021). How Should we Measure Public Sector Performance? *Postmodern Openings*, *1*. https://www.ceeol.com/search/article-detail?id=970053

OECD. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. https://www.oecd.org/sdd/42495745.pdf

OECD. (2020). *A Territorial Approach to the Sustainable Development Goals*. https://doi.org/10.1787/e86fa715-en

Ogata, K., Inoue, S., Ueda, A., & Yagi, H. (2018). The Functional Differentiation Between the International Integrated Reporting Council (IIRC) and the Global Reporting Initiative (GRI) in the Sphere of Sustainability Reporting. In K.-H. Lee & S. Schaltegger (Eds.), *Accounting for Sustainability: Asia Pacific Perspectives* (pp. 261–279). Springer International Publishing. https://doi.org/10.1007/978-3-319-70899-7_11

Olson, O., Humphrey, C., & Guthrie, J. (2001). Caught in an evaluatory trap: A dilemma for public services under NPFM. *European Accounting Review*, *10*(3), 505–522. https://doi.org/10.1080/09638180126799

Ortiz-Moya, F., Koike, H., Ota, J., Kataoka, Y., & Fujino, J. (2020). *Analysis of Voluntary Local Reviews* (STATE OF THE VOLUNTARY LOCAL REVIEWS 2020, pp. 14–29). Institute for Global Environmental Strategies; JSTOR. http://www.jstor.org/stable/resrep24977.12

Owen, D. L., Swift, T. A., Humphrey, C., & Bowerman, M. (2000). The new social audits: Accountability, managerial capture or the agenda of social champions? *The European Accounting Review*, *9*(1), 81–98. https://doi.org/10.1080/096381800407950

Pennington, L. K., & More, E. (2010). Sustainability Reporting: Rhetoric Versus Reality? *Employment Relations Record*, *10*(1), 24–39.

Pham, L. V., & Smith, C. (2013). Agricultural Sustainability in Developing Countries: An Assessment of the Relationships Between Drivers and Indicators in Hoa Binh Province, Vietnam. *Agroecology and Sustainable Food Systems*, *37*(10), 1144–1186. https://doi.org/10.1080/21683565.2013.833153

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

Rat für Nachhaltige Entwicklung. (n.a.). *German Council for Sustainable Development*. https://www.nachhaltigkeitsrat.de/en/

Reggi, L., Arduini, D., Biagetti, M., & Zanfei, A. (2014). How advanced are Italian regions in terms of public e-services? The construction of a composite indicator to analyze patterns of innovation diffusion in the public sector. *Telecommunications Policy*, *38*(5–6). https://doi.org/10.1016/j.telpol.2013.12.005

Regione Lombardia. (2022). *Voluntary Local Review 22: Lombardy Region Metropolitan City of Milan*.

RKA. (n.d.). *Kolada*. Retrieved May 23, 2023, from https://www.kolada.se/

Romer, P. M. (2015). Mathiness in the Theory of Economic Growth. *American Economic Review*, *105*(5). https://doi-org.kuleuven.e-bronnen.be/10.1257/aer.p20151066.

Roundy, P. T., Trussel, J. M., & Davenport, S. A. (2022). The text complexity of local government annual reports. *Local Government Studies*, *ahead-of-print*(ahead-of-print), 1–22. https://doi.org/10.1080/03003930.2022.2152009

Ryan, C., Dunstan, K., & Brown, J. (2002). The Value of Public Sector Annual Reports and Annual Reporting Awards in Organizational Legitimacy. *Accounting, Accountability and Performance*, *8*(1). https://www.proquest.com/scholarly-journals/value-public-sector-annual-reports-reporting/docview/220858394/se-2

Saldaña, Johnny. (2009). *The coding manual for qualitative researchers.* Sage Publications.

Samkin, G., & Schneider, A. (2010). Accountability, narrative reporting and legitimation. *Accounting, Auditing & Accountability Journal*, *23*(2), 256–289. https://doi.org/10.1108/09513571011023219

Saner, R., Yiu, L., & Nguyen, M. (2020). Monitoring the SDGs: Digital and social technologies to ensure citizen participation, inclusiveness and transparency. *Development Policy Review*, *38*(4), 483–500. https://doi.org/10.1111/dpr.12433

Scerri, A., & James, P. (2010). Accounting for sustainability: Combining qualitative and quantitative research in developing 'indicators' of sustainability. *International Journal of Social Research Methodology*, *13*(1), 41–53. https://doi.org/10.1080/13645570902864145

Schmidt-Traub, G., Kroll, C., Teksoz, K., Durand-Delacre, D., & Sachs, J. D. (2017). National baselines for the Sustainable Development Goals assessed in the SDG Index and Dashboards. *Nature Geoscience*, *10*. https://doi.org/10.1038/ngeo2985

SDSN Secretariat. (2014). *Framing Sustainable Development Goals, Targets, and Indicators*. Sustainable Development Solutions Network; JSTOR. http://www.jstor.org/stable/resrep16088

SDSN Youth. (n.d.). *What are Sustainable Development Goals (SDGs)?* Retrieved May 23, 2023, from https://sdgzone.com/learn/what-are-the-sdgs/

Shields, D. J., Šolar, S. V., & Martin, W. E. (2002). The role of values and objectives in communicating indicators of sustainability. *Ecological Indicators*, *2*(1), 149–160. https://doi.org/10.1016/S1470-160X(02)00042-0

Siragusa, A., Bertozzi, C., Stamos, I., & Proietti Paola. (2022). *European Handbook for SDG Voluntary Local Reviews—2022 Edition*. Publications Office of the European Union. 10.2760/218321

Spillane, J. P. (2012). Data in Practice: Conceptualizing the Data-Based Decision-Making Phenomena. *American Journal of Education*, *118*(2), 113–141.

Stadsledningskontoret. (2021). *Voluntary Local Review City of Stockholm 2021*.

State Capital Stuttgart. (2021). *The global 2030 Agenda at a local level: 2nd Voluntary Local Review*.

Steccolini, I. (2004). Is the Annual Report an Accountability Medium? An Empirical Investigation into Italian Local Governments. *Financial Accountability & Management*, *20*(3), 327–350. https://doi.org/10.1111/j.0267-4424.2004.00389.x

Stefanescua, A., Oprisor, T., & Sntejudeanu, M. A. (2016). An original assessment tool for transparency in the public sector based on the integrated reporting approach. *Accounting and Management Information Systems*, *15*(3), 542.

Steurer, R., & Hametner, M. (2013). Objectives and Indicators in Sustainable Development Strategies: Similarities and Variances across Europe: Objectives and Indicators in Sustainable Development Strategies. *Sustainable Development (Bradford, West Yorkshire, England)*, *21*(4), 224–241. https://doi.org/10.1002/sd.501

Swianiewicz, P. (2020). Recent and Contemporary Trends in European Studies of Local Government and Local Politics. In *Contemporary Trends in Local Governance* (pp. 21–43). https://doi-org.kuleuven.e-bronnen.be/10.1007/978-3-030-52516-3_2

The State Government of NRW. (2016). *Sustainability Strategy for North Rhine-Westphalia*.

Thiel, S. V., & Leeuw, F. L. (2002). The Performance Paradox in the Public Sector. *Public Performance & Management Review*, *25*(3), 267. https://doi.org/10.2307/3381236

Thomson, I. (2015). 'But does sustainability need capitalism or an integrated report' a commentary on 'The International Integrated Reporting Council: A story of failure' by Flower, J. *Critical Perspectives on Accounting*, *27*, 18–22. https://doi.org/10.1016/j.cpa.2014.07.003

Tiwari, G., Chauhan, S. S., & Varma, R. (2021). Challenges of localizing sustainable development goals in small cities: Research to action. *International Association of Traffic and Safety Sciences*, *45*(1), 3–11.

Tufte, E. R. (2001). The visual display of quantitative information. *Graphics Press*.

UNDESA. (2020). *Global Guiding Elements for Voluntary Local Reviews (VLRs) of SDG implementation*. https://sdgs.un.org/sites/default/files/2020-10/GlobalGuidingElementsforVLRs_FINAL.pdf

UNDP. (2023). *THE SDGS IN ACTION*. https://www.undp.org/sustainable-development-goals

UN-Habitat. (2018). *Leading Locally: The Origins and Impact of the Voluntary Local Review*. https://unhabitat.org/leading-locally-the-origins-and-impact-of-the-voluntary-local-review

UN-Habitat. (n.a.). *VOLUNTARY LOCAL REVIEWS*. https://unhabitat.org/topics/voluntary-local-reviews

UN-Habitat, UNDP, & UCLG. (2019). *SDGs Learning Module 3: Reporting to national & local reviews*. https://unhabitat.org/sites/default/files/2021/08/m3_en_web.pdf

United Cities and Local Governments. (2022). *Towards the Localization of the SDGs*. https://www.global-taskforce.org/sites/default/files/2022-07/Towards%20the%20Localization%20of%20the%20SDGs_2022.pdf

United Cities and Local Governments & UN-Habitat. (2020). *Guidelines for Voluntary Local Reviews: Volume 1*. https://unhabitat.org/guidelines-for-voluntary-local-reviews-volume-1-a-comparative-analysis-of-existing-vlrs

United Cities and Local Governments & UN-Habitat. (2021). *Guidelines for Voluntary Local Reviews: Volume 2*. https://unhabitat.org/sites/default/files/2021/07/vlrguidelines_vol2.pdf

United Nations. (2017a). *Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda*. https://unsdg.un.org/sites/default/files/UNDG_BigData_final_web.pdf

United Nations. (2017b). *Resolution adopted by the General Assembly on 6 July 2017* (A/RES/71/313). United Nations. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N17/207/63/PDF/N1720763.pdf?OpenElement

United Nations. (2022). *The Sustainable Development Goals Report 2022*. https://unstats.un.org/sdgs/report/2022/The-Sustainable-Development-Goals-Report-2022.pdf

United Nations. (n.a.-a). *Leave No One Behind*. https://unsdg.un.org/2030-agenda/universal-values/leave-no-one-behind

United Nations. (n.a.-b). *Sustainability*. https://www.un.org/en/academic-impact/sustainability

United Nations. (n.a.-c). *Sustainable Development Goals*. https://www.un.org/en/sustainable-development-goals

United Nations. (n.a.-d). *Transforming Our World: The 2030 Agenda for Sustainable Development*. https://sdgs.un.org/2030agenda

United Nations. (n.a.-e). *Voluntary common reporting guidelines for voluntary national reviews at the high-level political forum for sustainable development (HLPF)*. https://hlpf.un.org/sites/default/files/migrated/documents/17346Updated_Voluntary_Guidelines.pdf

United Nations. (n.a.-f). *Voluntary Local Reviews*. https://sdgs.un.org/topics/voluntary-local-reviews

Van de Walle, S. (2008). *Comparing the performance of national public sectors: Conceptual problems*.

Van de Walle, S., & Van Dooren, W. (2009). How is Information Used to Improve Performance in the Public Sector? Exploring the Dynamics of Performance Information. *Connecting Knowledge and Performance in Public Services: From Knowing to Doing*. https://doi.org/10.2139/ssrn.1491644

Van Der Knaap, P. (2006). Responsive Evaluation and Performance Management: Overcoming the Downsides of Policy Objectives and Performance Indicators. *Evaluation*, *12*(3), 278–293. https://doi.org/10.1177/1356389006069135

Van Dooren, W., Bouckaert, G., & Halligan, J. (2015). Performance as a result of production process. In *Performance Management in the Public Sector* (p. 21). https://www.routledge.com/Performance-Management-in-the-Public-Sector/Dooren-Bouckaert-Halligan/p/book/9780415738101

van Thiel, S. (2014). *Research Methods in Public Administration and Public Management: An Introduction*. Taylor & Francis. https://books.google.ee/books?id=RLwTAwAAQBAJ

Verhoest, K. (2017). *Control by inputs, results or markets? The control of public organisations and their performance in the New Public Management: Testing the assumptions*. Public Management Institute Katholieke Universiteit leuven. https://soc.kuleuven.be/io/pubpdf/IO02060003_koenspmiv1.pdf

Vikström, F. (n.d.). *A descriptive study of indicators and statistics used to enable comparisons and improved services in municipalities and county councils*. SKL International. http://projects.sklinternational.se/tuselog/files/2013/07/Statistics_-FINAL.pdf

Walker, R. G. (2002). Are Annual Reports of Government Agencies really "General Purpose" if They Do Not Include Performance Indicators? *Australian Accounting Review*, *12*(1). https://doi.org/10.1111/j.1835-2561.2002.tb00194.x

Wolcott, H. F. (2008). *Writing up qualitative research*. Sage publications.

Yang, K., & Holzer, M. (2006). The Performance-Trust Link: Implications for Performance Measurement. *Public Administration Review* , *66*(1), 114–126.

Zhang, J. (Yibo). (2020). The Impact of Vivid Graphical Presentation of Financial Information in Digital Annual Reports on Investors' Impressions of Management and Firm Performance. *The Journal of Information Systems*, *34*(3), 233–253. https://doi.org/10.2308/isys-52533

# Appendix

## Appendix 1 Summary of Qualities of Indicators

| NAO, HM Treasury, Cabinet Office, Audit Commission and the Office for National Statistics (2001) | Bird et al. (2005) | Bouckaert (1993) |
|---|---|---|
| Focused on the organization's aims and objectives; | Indicators should be directly relevant to PM's primary objective, or be an obviously adequate proxy measure | Countable |
| Appropriate to, and useful for, the stakeholders that are likely to use it; | Definitions need to be price but practicable | Uniform over time, timeliness |
| Balanced, giving a picture of what the organization is doing, covering all significant areas of work; | Survey-based indicators should use a shared methodology and common questions between institutions. | Data readily available |
| Robust in order to withstand organizational changes or individuals leaving; | Indicators and definitions should be consistent over time. If, exceptionally, changes need to be made, these should be documented in a PM and an impact assessment made. Otherwise, a new PM base-line needs to be set. | Accuracy |
| Integrated into the organization, being part of the business planning and management process | Indicators and definitions should obviate, rather than create, perverse behaviors | Understandable, unequivocal |
| Cost effective, balancing the benefits of the information against the costs | Indicators should be straightforward to interpret, avoiding ambiguity about whether the performance being monitored has improved or deteriorated. Ambiguity arises if there is major unaccounted-for disparity in case mix of individuals received into institutions, or if case mix (of pupils entered for examinations, patients selected for operation, or offend- ers given police cautions (House of Commons Public Administration Select Committee, 2003d)) drifts over time | Comprehensiveness |
| | Indicators that are not collected for the whole population should have sufficient cover- age to ensure against misleading results, i.e. potential bias compared with measuring the target population should be small. | Congruent |
| | Technical properties of the indicator should be adequate, e.g. with respect scheme, response rate in surveys, and precision (standard error of difference | Reproducible |
| | Indicators should have the statistical potential to exhibit or identify change intended timescale of PM | Objective |
| | Indicators should be produced with appropriate frequency, disaggregation for context, and timeliness to support performance management | Choosable |

| Better Regulation guidelines of the European Commission |
| --- |
| Relevant, i.e. closely linked to the objectives to be reached<br><br>Accepted (e.g. by staff, stakeholders). The role and responsibilities of the indicator need to be well defined;<br><br>Credible for non-experts, unambiguous and easy to interpret. Indicators should be as simple and robust as possible<br><br>Easy to monitor (e.g. data collection should be possible at low cost)<br><br>Robust against manipulation (e.g. administrative burden: if the target is to reduce administrative burdens to |

## Appendix 2 IIRF Guideline

| Fundamental Concepts | Guiding Principles | Content Elements |
|---|---|---|
| Value creation for the organization and others | Strategic focus and future orientation: An integrated report should provide insight into the organization's strategy, and how it relates to the organization's ability to create value in the short, medium and long term, and to its use of and effects on the capitals | Organizational overview and exter environment: What does the organization do and what are the circumstances under which it operates? |
| The capitals | Connectivity of information: An integrated report should show a holistic picture of the combination, interrelatedness and dependencies between the factors that affect the organization's ability to create value over time | Governance: How does the organization's governance structure support its ability to create value in the short, medium and long term? |
| | Stakeholder relationships: An integrated report should provide insight into the nature and quality of the organization's relationships with its key stakeholders, including how and to what extent the organization understands, takes into account and responds to their legitimate needs and interests | Business model: What is the organization's business model? |
| | Materiality: An integrated report should disclose information about matters that substantively affect the organization's ability to create value over the short, medium and long term | Risks and opportunities: What are the specific risks and opportunities that affect the organization's ability to create value over the short, medium and long term, and how is the organization dealing with them? |
| | Reliability and completeness: An integrated report should include all material matters, both positive and negative, in a balanced way and without material error | Strategy and resource allocation: Where does the organization want to go and how does it intend to get there? |
| | Consistency and comparability: The information in an integrated report should be presented: (a) on a basis that is consistent over time; and (b) in a way that enables comparison with other organizations to the extent it is material to the organization's own ability to create value over time. | Performance: To what extent has the organization achieved its strategic objectives for the period and what are its outcomes in terms of effects on the capitals? |
| | Conciseness: An integrated report should be concise | Outlook: What challenges and uncertainties is the organization likely to encounter in pursuing its strategy, and what are the potential implications for its business model and future performance? |
| | | Basis of preparation and presentation: How does the organization determine what matters to include in the integrated report and how are such matters quantified or evaluated? |

## Appendix 3 GRI Guideline

| Key Concepts | Reporting Principles | Requirements |
|---|---|---|
| Impact: the effect an organization has or could have on the economy, environment, and people, including effects on their human rights, as a result of the organization's activities or business relationships. | Accuracy | 1: Apply the reporting principles |
| Material topic: areas of impact | Balance | 2: Report the disclosures in GRI 2: General Disclosures 2021 |
| Due diligence: the process through which an organization identifies, prevents, mitigates, and accounts for how it addresses its actual and potential negative impacts on the economy, environment, and people, including impacts on their human rights | Clarity | 3: Determine material topics |

| | | |
|---|---|---|
| Stakeholder: individuals or groups that have interests that are affected or could be affected by an organization's activities. | Comparability | 4: Report the disclosures in GRI 3: Material Topics 2021 |
| | Completeness | 5: Report disclosures from the GRI Topic Standards for each material topic |
| | Sustainability context | 6: Provide reasons for omission for disclosures and requirements that the organization cannot comply with |
| | Timeliness | 7: Publish a GRI content index |
| | Verifiability | 8: Provide a statement of use 9: Notify GRI |

## Appendix 4 Requirement 1 GRI Reporting Principles and Guidance

| Reporting Principles | Guidance |
|---|---|
| Accuracy | • report qualitative information that is consistent with available evidence and other reported information<br>• indicate which data has been measured; adequately describe data measurements and bases for calculations, and ensure it is possible to replicate measurements and calculations with similar results<br>• ensure that the margin of error for data measurements does not inappropriately influence the conclusions or assessments of information users<br>• indicate which data has been estimated, and explain the underlying assumptions and techniques used for the estimation as well as any limitations of the estimates |
| Balance | • present information in a way that allows information users to see negative and positive year-on-year trends in impacts<br>• distinguish clearly between facts and the organization's interpretation of the facts<br>• not omit relevant information concerning its negative impacts<br>• not overemphasize positive news or impacts<br>• not present information in a way that is likely to inappropriately influence the conclusions or assessments of information users |
| Clarity | • consider specific accessibility needs of information users, associated with abilities, language, and technology<br>• present information in a way that users can find the information they want without unreasonable effort, for example, through a table of contents, maps, or links<br>• present information in a way that it can be understood by users who have reasonable knowledge of the organization and its activities<br>• avoid abbreviations, technical terms, or other jargon likely to be unfamiliar to users or, if these are used, include relevant explanations in the appropriate sections or in a glossary<br>• report information in a concise way and aggregate information where useful without omitting necessary details<br>• use graphics and consolidated data tables to make information accessible and understandable |
| Compara-bility | • present information for the current reporting period and at least two previous periods, as well as any goals and targets that have been set<br>• use accepted international metrics (e.g., kilograms, liters), and standard conversion factors and protocols, where applicable, for compiling and reporting information<br>• maintain consistency in the methods used to measure and calculate data and in explaining the methods and assumptions used<br>• maintain consistency in the manner of presenting the information<br>• report total numbers or absolute data (e.g., metric tons of CO2 equivalent) as well as ratios or normalized data (e.g., CO2 emissions per unit produced) to enable comparisons, and provide explanatory notes when using ratios<br>• provide contextual information (e.g., the organization's size, geographic location) to help information users understand the factors that contribute to differences between the organization's impacts and the impacts of other organizations<br>• present the current disclosures alongside restatements of historical data to enable comparisons if there have been changes from the information reported previously. This can include changes in the length of the reporting period, in the measurement methodologies, in the definitions used, or in other elements of reporting. The organization is required to report restatements of information under Disclosure 2-4 in GRI 2: General Disclosures 2021 |

| | |
|---|---|
| | • if restatements of historical data are not provided, explain the changes to provide contextual information for interpreting the current disclosures |
| Complete-ness | • present activities, events, and impacts for the reporting period in which they occur. This includes reporting information about activities that have a minimal impact in the short-term, but a reasonably foreseeable cumulative impact that can become unavoidable or irreversible in the long-term (e.g., activities that generate bio-accumulative or persistent pollutants)<br>• not omit information that is necessary for understanding the organization's impacts |
| Sustain-ability context | • draw on objective information and authoritative measures on sustainable development to report information about its impacts (e.g., scientific research or consensus on the limits and demands placed on environmental resources)<br>• report information about its impacts in relation to sustainable development goals and conditions (e.g., reporting total greenhouse gas [GHG] emissions as well as reductions in GHG emissions in relation to the goals set out in the United Nations [UN] Framework Convention on Climate Change [FCCC] Paris Agreement [4])<br>• report information about its impacts in relation to societal expectations and expectations of responsible business conduct set out in authoritative intergovernmental instruments with which the organization is expected to comply (e.g., Organisation for Economic Co-operation and Development [OECD] Guidelines for Multinational Enterprises [3], UN Guiding Principles on Business and Human Rights [5]) and in other recognized sector-specific, local, regional, or global instruments<br>• if operating in a range of locations, report information about its impacts in relation to appropriate local contexts (e.g., reporting total water use, as well as water use relative to the sustainable thresholds and the social context of given catchments) |
| Timeliness | • find a balance between the need to make information available in a timely manner and ensuring that the information is of high quality and meets the requirements under the other reporting principles<br>• ensure consistency in the length of reporting periods<br>• indicate the time period covered by the reported information |
| Verifi-ability | • set up internal controls and organize documentation in such a way that individuals other than those preparing the reported information (e.g., internal auditors, external assurance providers) can review them<br>• document the decision-making processes underlying the organization's sustainability reporting in a way that allows for the examination of the key decisions and processes, such as the process of determining material topics<br>• if the organization designs information systems for its sustainability reporting, design these systems in a way that they can be examined as part of an external assurance process<br>• be able to identify the original sources of the reported information and provide reliable evidence to support assumptions or calculations<br>• be able to provide representation from the original sources of the reported information attesting to the accuracy of the information within acceptable margins of error<br>• avoid including information that is not substantiated by evidence unless it is relevant for understanding the organization's impacts<br>• provide clear explanations of any uncertainties associated with the reported information |

## Appendix 5 Introduction of the SDGs and Mapping to Local Context

| LRG | SDGs/Dimensions | Mapping to local context |
|---|---|---|
| City of Ghent | SDGs | Goal |
| Municipality of Gladsaxe | Dimensions | Groups of goals |
| City of Tampere | SDGs | Goal |
| City of Helsinki | Dimensions | Goal |
| City of Vantaa | SDGs | Goal |
| City of Espoo | Dimensions | Goal |
| City of Turku | SDGs | Goal |
| City of Bonn | Dimensions | Target |
| City of Düsseldorf | Dimensions | Target |
| City of Kiel | HLPF SDGs | Target |
| City of Stuttgart | SDGs | Target |
| City of Hanover | Dimensions | Target |
| City of Mannheim | Dimensions | Goal |
| State of North Rhine-Westphalia | Dimensions | Goal |
| City of Skiathos | Dimensions | Goal |
| Region of Lombardy | SDGs | Goal |
| City of Amsterdam | SDGs | Goal |
| City of Barcelona | SDGs | Target |
| Basque Country Government | SDGs | Target |
| The Generaliat Valencia | None | None |
| City of Stockholm | SDGs | Goals |
| City of Malmo | HLPF SDGs | Target |
| City of Helsingborg | SDGs | Goals |

## Appendix 6 Summary of objectives

| LRG | **Account** | **Learn** | **Steer & control** | **Communicate** |
|---|---|---|---|---|
| City of Ghent | x | x | | |
| Municipality of Gladsaxe | x | x | | |
| City of Tampere | x | x | x | |
| City of Helsinki | x | x | x | |
| City of Vantaa | | x | | |
| City of Espoo | x | | | |
| City of Turku | | x | x | x |
| City of Bonn | x | x | | |
| City of Düsseldorf | x | x | | |
| City of Kiel | | | | x |
| City of Stuttgart | | x | | |
| City of Hanover | x | | | |
| City of Mannheim | | | | |

| | | | | |
|---|---|---|---|---|
| State of North Rhine-Westphalia | x | | | |
| City of Skiathos | x | | x | |
| Region of Lombardy | | | | |
| City of Amsterdam | x | | | |
| City of Barcelona | | x | x | |
| Basque Country Government | x | | | |
| The Generaliat Valencia | | | | |
| City of Stockholm | x | | | |
| City of Malmo | x | x | x | |
| City of Helsingborg | | | x | |

## Appendix 7 Details of data sources

| LRG | Data source | Internal/external |
|---|---|---|
| City of Ghent | OECD | External |
| | Datacel Social Services City of Ghent, 2020 | Internal |
| | UAntwerpen & UHasselt &CU Leuven | External |
| | Sixth COVID-19 Health Survey. First results. Brussels, Belgium | External |
| | Education Services Agency | External |
| | own figures urban education | Internal |
| | City of Ghent Childcare Department | Internal |
| | City of Ghent, Population statistics 2020. | Internal |
| Municipality of Gladsaxe | N/A | |
| City of Tampere | Business Tampere | Internal |
| | City of Tampere | Internal |
| | CO2 report (Sitowise) | External |
| | FinChildren register monitoring (THL) | External |
| | FinSote survey (THL) | External |
| | Ministry of Employment and the Economy | External |
| | Pirkanmaan Voimia | Internal |
| | Police | External |
| | School Health Promotion study (THL) | External |
| | Statistics Finland | External |
| | Taloustutkimus | External |
| | Tampere Regional Solid Waste Management Ltd | External |
| | Tampere Water | Internal |
| | Tampereen Sahkolaitos | Eternal |
| | Well-being survey of Tampere | Internal |

| | | |
|---|---|---|
| City of Helsinki | Health and Welfare Barometer | Internal |
| | Helsinki Welfare Report 2018-2021 | Internal |
| | Statistic Finland | External |
| City of Vantaa | N/A | |
| City of Espoo | Kela | External |
| | espoo.fi | Internal |
| | fortum.com/espoo | External |
| | Caruna and the City of Espoo | External |
| | Urban and municipal services survey 2019 | Internal |
| | HSY.fi | External |
| | the City of Espoo, SYKE and MML | External |
| | Natural Resources Insititute Finland. | External |
| | Statics Finland and the Ministry of Justice | External |
| City of Turku | Sotkanet, THL | External |
| | Ara: Asunnottomat 2018 | External |
| | National study on passenger traffic, WSP Finland Oy | External |
| | Hyvinvointikompassi, THL | External |
| City of Bonn | N/A | |
| City of Düsseldorf | SDG Portal | External |
| | State Capital Dusseldorf, Office for the Environment and Consumer Protection | Internal |
| | IT.NRW, area survey by type of actual use | Internal |
| | State Capital Dusseldorf, Office for Statistics and Elections | Internal |
| | State Capital Dusseldorf, Garden, Cemetry and Forestry Office, Lower Nature Conservation Authority | Internal |
| | SDG Portal, *Source for 2010: City of Dusseldorf, Canal and Water Engineering Office | |
| | TU Dresden, SrV - Mobili- ty in cities | External |
| | Federal Motor Transport Authority | External |
| | State Capital Dusseldorf, Office for Traffic Management | Internal |
| | State Capital Dusseldorf, Youth Welfare Office | Internal |
| | Destatis | External |
| | IT.NRW, VGRdL | Internal |
| | Statistics of the Federal Employment Agency | External |
| | www.fairtrade- schools.de | External |
| City of Kiel | SDG Portal | External |
| | https://www.kiel.de/de/kiel_zukunft/ nachhaltigkeitsziele/hochwertige_bildung.php | Internal |
| | Day-care reform law in Schleswig-Holstein | External |
| | Port of Kiel Waste Management Plan 2019/2020 | Interla |
| | Office of the City Council President, International Affairs and Sustainability, City of Kiel | Internal |

| | | |
|---|---|---|
| | Christian-Albrecht University of Kiel, Kiel University of Applied Sciences and Muthesius University of Fine Arts and Design | External |
| City of Stuttgart | Federal Labor Office, State Capital Stuttgart | Internal |
| | "Wegweiser Kommune" of the Bertelsmann Foundation, State Capital Stuttgart, Statistics Office | Internal |
| | State Capital Stutttgart, City Treasury | Internal |
| | State Statistical Office (Agricultural structure survey) | Internal |
| | Bertelsmann Foundation, University of Giessen, Institute for Resource Management | External |
| | State Capital Stuttgart, Office of Sport and Physical Activity | Internal |
| | State Statistics Office of Baden-Württemberg | External |
| | Federal Office for Building and Regional Planning / Federal Pharmacies' Registry | External |
| | State Capital Stuttgart, Environmental Protection Office | Internal |
| | State Statistical Office; State Capital Stuttgart, Youth Welfare Office and Statistics Office | Internal |
| | State Capital Stuttgart, Environmental Protection Office, Parks, Cemeteries and Forestry Office; AWS; Stadtwerke Stuttgart | Internal |
| | State Capital Stuttgart, Library of Stutttgart | Internal |
| | State Capital Stuttgart, Cultural Affairs Office | Internal |
| | City Surveying Office as of 2020, brochure "Bildung natürlich" Stuttgart (2020), www.stuttgart.de/handlungsfelder-bildungsgerechtigkeit, Keyword "Bildung natürlich" | Internal |
| | "National Accounts of the Länder" working group | External |
| | Federal and State Statistical Offices; Federal Labour Office; State Capital Stuttgart, Statistics Office | External/internal |
| | Bertelsmann Foundation, Wegweiser Kommune | External |
| | State Capital Stuttgart, Statistics Office (immoscout24.de) | Internal |
| | State Capital Stuttgart, Urban Planning and Housing Office | Internal |
| | State Capital Stuttgart, Policy Planning and Sustainable Mobility Division | Internal |
| | State Capital Stuttgart, Public Undertaking Stuttgart Waste Management (AWS) | Internal |
| | German Chamber of Industry and Commerce, evaluation by The German Institute for Urban Studies | External |
| City of Hanover | Bertelsmann Stiftung "Monitor Sustainable Municipality", Report 2016, Part 1, p. 19 | External |
| | City of Hanover, Department of the Mayor's Office, policy matters | Internal |
| | Engagement Global (2019): Sustainable Development Goals [20] | External |
| | City of Hanover, Environment and Urban Greenspace Division | Internal |
| | Statistical Office of the Federal State of Lower Saxony, Statistical Yearbook 2019 | External |
| | Region of Hanover, 2015: energy and CO2 balance for the Region of Hanover | Internal |
| | Region of Hanover, 2019: energy and CO2 balance for the Region of Hanover and own calculations by the City of Hanover's control centre for climate change mitigation | Internal |
| | City of Hanover, Facility Management Division | Internal |
| | GEO-NET (2016): Analysis of the climatic and ecological functions and processes for the City of Hanover | External |
| | Federal Ministry of Transport and Digital Infrastructure, 2018: Mobilität in Deutschland 2017; Region of Hanover | External |
| | ÜSTRA Hannoversche Verkehrsbetriebe AG | External |
| | City of Hanover, Public Order Department | Internal |

| | | |
|---|---|---|
| | Annual reports on air quality, state trade supervisory office of Hildesheim | Internal |
| | Institut für Umwelt-Analyse Projekt GmbH (IFUA), 2009 and 2019; City of Hanover, Environment and Urban Greenspace Division | External |
| | enercity AG, Hanover | External |
| | Arbeitsgemeinschaft Limnologie und Gewässerschutz e. V., Hanover | External |
| | Zweckverband Abfallwirtschaft Region Hannover (aha) | External |
| | City of Hanover, Agenda 21 and Sustainability Office | Internal |
| | Federal Employment Agency \| From: Statistical Yearbook of the City of Hanover 2020, Elections and Statistics Department | External |
| | Lower Saxony state office for statistics<br>From: City of Hanover, Elections and Statistics Department | External |
| | Lower Saxony state office for statistics | External |
| | empirica price database; analysis: City of Hanover, Urban Development | External |
| | Statistics from the Federal Employment Agency and the City of Hanover, Elections and Statistics Department | External |
| | City of Hanover, Social Affairs Division as well as Elections and Statistics Department and the Federal Employment Agency | Internal |
| | Stadtsportbund Hannover e. V.; Statistical Yearbook of the City of Hanover 2020, Elections and Statistics Department | Internal |
| City of Mannheim | N/A | |
| State of North Rhine-Westphalia | N/A | |
| City of Skiathos | N/A | |
| Region of Lombardy | N/A | |
| City of Amsterdam | CBS, Health Monitor 2020. | External |
| | Kennisplatform Inclusief Samenleven, KIS-wijkmonitor, 2021. | External |
| | Unie van Waterschappen | External |
| | Waar staat je gemeente, 2022 | External |
| | Staat van de Stad 2021, p. 49 | Internal |
| | Factsheet Schulden in Amsterdam 2021 Gemeente Amsterdam. | Internal |
| | MRA, Economische Verkenningen Metropoolregio Amsterdam 2021, p. 82-84. | External |
| | Jobdigger / OIS. | External |
| | CBS, Arbeidsdeelname; regionale indeling 2020. | External |
| | Kennisplatform Inclusief Samenleven (2021). | External |
| | Drechtsteden Research Center (no date), Municipal Sustainability Index. | External |
| | Vluchtelingenmonitor 2020. | External |
| | Migratiemonitor Metropoolregio (2020), p. 16. | External |
| | Wonen in Amsterdam, 2021. | Iinternal |
| | Actieplan Schone Lucht (2019). | Internal |
| | maps.amsterdam.nl/geluid/. | Internal |
| | UN | External |
| City of Barcelona | | |
| Basque Country Government | Basque Sociometer, Office of Sociological Research, Presidency of the Government | Internal |

| | | |
|---|---|---|
| The Generaliat Valencia | | |
| City of Stockholm | The national public health survey (FHE). | External |
| | Sweco | External |
| | The National Education Agency. | External |
| | Source: The Swedish Election Authority, Sweco. | External |
| | Environment and Health Administration | External |
| | Environmentally-friendly vehicles and renewable fuels in Stockholm, ENTSO-E, Eurostat and Stockholm Energi. | External |
| | Miljobarometern.se, the City of Stockholm. | Internal |
| | SL and the city's own traffic surveys. | Internal |
| | The City of Stockholm's planning department. | Internal |
| | SL and the city's own traffic surveys. | Internal |
| | Miljobarometern.se, the City of Stockholm. | Internal |
| | Kolada | External |
| | Environment Programme 2020–2023, the City of Stockholm. | Internal |
| | Hantera livs (Climate data base for food). | External |
| City of Malmo | Statistics Sweden | External |
| | Swedish Board of Health and Welfare (kolada ID: 31816) | External |
| | Swedish Public Health Agency, Kolada ID (U01411) | External |
| | Research study Young Lifestyle (Ung livsstil 2018) | External |
| | Swedish Social Insurance Agency Kolada ID (N74811) | External |
| | Swedish Municipalities and Region's worker and salary statistics Kolada ID (N00209) | External |
| | City of Helsingborg | Internal |
| | Öresundskraft AB | External |

# Appendix 8 Type and number of visualizations

| LRG | Total number of visualizations | Line | Line plus bar | Horizontal bar | Vertical bar | Pie | Spider | Map | Picto | Text |
|---|---|---|---|---|---|---|---|---|---|---|
| Average | 2.3 | | | | | | | | | |
| City of Ghent | 4 | x | x | | x | | | x | | x |
| Municipality of Gladsaxe | 0 | | | | | | | | | x |
| City of Tampere | 0 | | | | | | | | | x |
| City of Helsinki | 4 | x | | | x | | x | x | | x |
| City of Vantaa | 3 | | | x | | x | | x | | |
| City of Espoo | 7 | x | x | x | x | x | | x | x | |
| City of Turku | 1 | | | | | | | | x | |
| City of Bonn | 3 | x | | x | x | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| City of Düsseldorf | 3 | x | | x | x | | | | | | |
| City of Kiel | 1 | | | | | | | | x | | |
| Stuttgart | 3 | x | | | x | x | | | | | |
| City of Hanover | 4 | x | x | | x | x | | | | | x |
| City of Mannheim | 0 | | | | | | | | | | |
| State of North Rhine-Westphalia | 0 | | | | | | | | | | |
| City of Skiathos | 2 | | | | | | | | | x | x |
| Region of Lombardy | 3 | x | | | x | | x | | | | x |
| City of Amsterdam | 5 | x | | | x | x | | x | x | | |
| Barcelona | 1 | | x | | | | | | | | |
| Basque Country Government | 0 | | | | | | | | | | x |
| The Generaliat Valencia | 0 | | | | | | | | | | |
| City of Stockholm | 2 | x | | | x | | | | | | x |
| City of Malmo | 4 | x | x | | x | | | x | | | |
| City of Helsingborg | 3 | x | | | x | x | | | | | |

## Appendix 9 Calculation of correlation between leading and lagging indicators

| | *Leading* | *Lagging* |
|---|---|---|
| Leading | 1 | |
| Lagging | 0.623638447 | 1 |