Urban Planning
2025 • Volume 10 • Article 10266
https://doi.org/10.17645/up.10266

cogitatio

ARTICLE

Open Access Journal

Rethinking Urban Sustainability Monitoring: Lessons Learnt From the SDG Indicators for Municipalities in Germany

Oliver Peters [1,2] and Valeska Liedloff [2]

[1] Institute of Urban Planning, Brandenburg University of Technology Cottbus-Senftenberg, Germany
[2] German Institute of Urban Affairs (Difu), Germany

Correspondence: Oliver Peters (opeters@difu.de)

Submitted: 3 March 2025 **Accepted:** 30 September 2025 **Published:** 10 November 2025

Issue: This article is part of the issue "Future Urban Sustainability: Lessons Learnt From the SDGs and Perspectives for a Post-2030 Agenda" edited by Florian Koch (HTW Berlin), Sarah Beyer (HTW Berlin), Kerstin Krellenberg (University of Vienna), and Julia Wesely (University of Vienna), fully open access at https://doi.org/10.17645/up.i433

Abstract

The translation of global sustainability agendas into urban strategies remains a challenge for local governments. Nevertheless, the increasing number of municipalities establishing municipal sustainability management systems and publishing sustainability reports indicates a growing commitment by cities and regions to achieve the Sustainable Development Goals (SDGs). It also points to improvements in the way inputs, outputs, outcomes, and impacts are measured at the local level. This article draws on seven years of experience with the well-established SDG Indicators for Municipalities in Germany to provide a critical examination of the development and application of sustainable urban development indicators. As significant methodological and data deficits persist, fundamental questions arise for future monitoring approaches—for instance, regarding the necessity of quantifying sustainability aspects and the suitability of the SDGs for local action and monitoring. In advance of a post-2030 agenda, the utility of the existing SDG framework is increasingly being reassessed due to the existence of thematic gaps regarding sustainable urban development and normative trade-offs, including those related to social cohesion, basic public services, and the growth paradigm. SDG indicators at the local level naturally reproduce these limitations, unlike adapted monitoring systems that are tailored to local requirements, such as spatial and temporal granularity. Based on theoretical indicator evaluation frameworks and insights from municipal practice, this article identifies seven criteria that should be considered and balanced for the local monitoring of a future sustainability agenda.

Keywords

2030 Agenda; monitoring; sustainability indicators; sustainable urban development

1. Introduction

The 2030 Agenda finally paved the way for the widespread use of indicators as the commonly accepted monitoring mechanism for sustainable development. A universal and extensive roadmap such as the Sustainable Development Goals (SDGs), containing a wide range of targets that interact with each other, is reliant on evidence-driven tools for understanding progress and measuring success. However, according to Webb and Ayyub (2017), quantifying sustainability involves translating complex socio-ecological relationships into measurable indicators, and this presents challenges, especially in the area of sustainable urban development (SUD). SUD can be defined as follows:

> The spatial manifestation of urban development processes that creates a built environment with norms, institutions and governance systems enabling individuals, households and societies to maximize their potential, optimize a vast range of services so that homes and dynamic neighborhoods, cities and towns are planned, built, renewed and consolidated restraining adverse impacts on the environment while safeguarding the quality of life, needs and livelihood of its present and future populations. (UN-Habitat, 2012, p. 1)

Despite its many aspects, this comprehensive definition makes it clear that impact assessment plays a central role in the facilitation and optimisation of SUD processes. Against this background, indicators can fulfil various functions: Firstly, they are essential for enabling evidence-based policy management by measuring the effectiveness of policies and interventions over time and tracking progress toward sustainability goals. Secondly, sustainability indicators can be used to visualise changes over a certain period of time and therefore, thirdly, to promote communication and transparency within the municipal administration and towards the public (e.g., Giles-Corti et al., 2020; Webb & Ayyub, 2017).

Depending on how indicators are perceived, developed, and used, their effects on SUD can vary to a great extent. Generally, indicators can be understood in two main ways: as neutral, technical tools that provide measurable data for policy-making, and as "message carriers" that reflect political and normative interpretations of sustainability. This means that developing indicators is not only a technical task, but also a dynamic and political process that helps define what urban sustainability actually means (Merino-Saum et al., 2020). The development of this understanding can be clearly observed in practice in the process of creating sustainability reports, such as Voluntary Local Reviews, which widely deploy quantitative assessments. Ortiz-Moya and Reggiani (2023) have identified three main aspects in which Voluntary Local Review processes can contribute to SUD policy integration: by promoting collaboration and interconnections across various policy sectors, by developing new tools to integrate the SDGs into mainstream practices, and by strengthening sustainability competencies. In this respect, reporting itself, with its deliberative processes, can lead to better indicators and vice versa. However, Koch et al. (2023), among others, point out that improving indicators, data, and communication tools does not automatically lead to better SUD governance and thus to positive outcomes. Nevertheless, understanding the differences in sustainable development via indicators can also enhance the ability of central authorities to balance aspects between national and various subnational levels (Benedek et al., 2021).

In addition to the overarching question of what role indicators play in SUD, there are specific methodological questions that arise when establishing an indicator system—both globally and locally. First of

all, it should be noted that there is no shortage of indicators and measurement concepts. In fact, the opposite is true: A broad range of measurement initiatives have been developed for monitoring and comparing the sustainability performance of cities worldwide, with some authors already stating that there has been an explosion of indicators, driven by the vague concept of sustainability and increasing data availability (Merino-Saum et al., 2020). This abundance of indicators makes selection all the more important. However, no single and universally legitimised method for choosing indicators exists, and each evaluation method has its own advantages and disadvantages, which emphasises the need for guidance that can provide practitioners with science-based and informed support (Gebara et al., 2024; Shi et al., 2019; Stitteneder, 2025).

According to Tran (2016), the evaluation and selection of SUD indicators can be observed in two main processes: The top-down approach involves experts and researchers defining the overall sustainability framework and then identifying specific indicators within that structure. In contrast, the bottom-up approach emphasises the active involvement of various stakeholders to collaboratively shape the framework and identify key indicators. Although the distinction between these two approaches has become less clear recently, Merino-Saum et al. (2020) argue that both approaches can reinforce inherent tensions between competing goals and methodological principles. Some of these tensions relate to conflicting quality criteria for individual indicators, whereas others emerge when considering the indicator set in its entirety. In their study, they focus on the three trade-offs most frequently discussed in the literature: parsimony versus comprehensiveness, context-specificity versus general comparability, and complexity versus simplicity.

In Germany, the SUD indicator set "SDG Indicators for Municipalities" is the result of those methodological discussions (Jossin & Peters, 2022). Developed within the eponymous project, which has been carried out since 2018 with municipal stakeholders in Germany, the SDG Indicators for Municipalities aim at translating and adapting indicators to the local level for mapping the SDGs in German cities, counties, and towns. The indicator set is continuously developed through participatory processes that include regular group discussions with municipalities and working group sessions involving 10 key institutions—among them research organisations such as the German Institute of Urban Affairs, local government associations such as the German Association of Cities, foundations such as the Bertelsmann Stiftung, and government advisory bodies such as the German Council for Sustainable Development. With its transparent methodology and the provision of available indicator data in the associated SDG portal (https://sdg-portal.de/en), the indicator set now serves as a key resource for German municipalities (e.g., for the preparation of sustainability reports). The fourth indicator catalogue from 2024 contains 100 indicators to map around 75 percent of the targets of the 2030 Agenda—those that have been assessed as relevant for German municipalities from a scientific perspective (Bertelsmann Stiftung et al., 2024). The indicators originate from various sources, including global, European, and national catalogues, municipal reports, and information provided by many partners and project users. Other indicators are original project developments. Thus, the SDG Indicators for Municipalities can be seen as the result of a participatory top-down approach that—just like many other SUD indicator sets—must meet the needs of many scientific criteria and practical requirements in order to provide a low-threshold entry point into municipal sustainability management and achieve broad acceptance.

Since studies on indicator evaluation criteria for SUD rarely address expert and practitioner perspectives at the same time, and generally pay little attention to trade-offs within those criteria, this article intends to contribute to important questions on SDG localisation:

RQ1: What are the most relevant requirements for SUD indicator systems from both a scientific and a practical perspective?

RQ2: What interactions, synergies, and trade-offs can be identified between these requirements?

RQ3: In order to address the aforementioned requirements and interactions, what design options are available for indicator systems that facilitate their local application and therefore should be considered in a post-2030 agenda?

Answers to these research questions could support the development of future indicator systems on various governance levels. Key indicator evaluation criteria from relevant frameworks allow scholars and policy-makers to ensure the actual implementation of indicator systems, while enabling transparency and comparability. Practitioners, such as local authorities and relevant organisations, in turn could use those criteria as guidance to adapt general indicator systems by selecting, developing, and modifying indicators according to the local situation.

In order to best support these objectives, the article is structured as follows. Section 2 presents the methodological approach followed to collect and characterise the indicator evaluation criteria finally included in the analysis. Section 3 elucidates the study's results by presenting the reasoned selection of important criteria and providing individual descriptions and discussions. For each criterion, the analysis provides the identified theoretical criteria and tensions with other criteria, the practical requirements, and the derived future design options to handle tensions between different criteria and perspectives. Section 4 provides an overview of possible limitations, and Section 5 concludes the findings.

2. Methodology

To capture both the current scientific and practical discourse, the study is based on a mixed-method approach and two main data sources. For information on the practical applicability of criteria in the individual evaluation and selection of indicators on site, primary data from group discussions were collected. Since feedback from municipal representatives may not cover the scientific discourse on potential criteria, and since trade-offs and design options may be subject to local contextual factors or may already be the result of deliberative processes, this was preceded by extensive literature research.

The literature review was aimed at comparing and synthesising existing indicator evaluation frameworks (IEFs) with a special emphasis on SUD. Several well-known criteria sets and principles are employed by science, organisations, and practitioners to guide the development of monitoring systems and the selection of appropriate indicators. Therefore, IEFs were primarily extracted from practical guidelines for indicator development or standardised indicator systems of national or international bodies, which are usually published as grey literature, and supplemented by literature searches via Google Scholar and Web of Science. The literature review was conducted in November and December 2024 using a combination of different keywords, i.e., "indicator," "monitoring," "evaluation," "selection," "assessment," "criteria," "framework," "standard," "sustainable urban development," "sustainability," "local," "regional," "urban," and "cities." The identified criteria for SUD indicators, which are intended to show complex cause-and-effect relationships in simple key figures, are diverse. Accordingly, a variety of criteria grids have been established,

which have found their way into the field of indicator evaluation as acronyms, such as RACER, which stands for "Relevant, Acceptable, Credible, Easy and Robust" (European Commission, 2023).

The SDG Indicators for Municipalities play a special role in this analysis, as the underlying project provides both its own IEF and access to the practical requirements from municipalities, which the authors of this article collected in group discussions in their role as responsible researchers for the indicator development. The SDG Indicators for Municipalities project was launched in 2017 and is ongoing, i.e., the indicators are continuously being developed, which usually leads to a new edition of the indicators every two years. On the one hand, this allows incorporating new developments from all the indicator catalogues on which the SDG Indicators for Municipalities are based, and which are then developed further simultaneously. On the other hand, general and specific methodological, topical, and data deficits also persist (Jossin & Peters, 2022; Mori & Christodoulou, 2012; Nagy et al., 2018; Verma & Raghubanshi, 2018). Many aspects can be explained by the translation of an abstract global target system into a local indicator system, which is intended to provide concrete governance. In the process, information can be lost or unintended assumptions can be made that call the validity of the indicators into question. Therefore, continuous reflection on the criteria and their application is crucial to the project—this participatory approach forms the second source of the present analysis.

A total of five central workshops with representatives from German cities, counties, and towns and partner organisations were conducted during the phases of refining the SDG Indicators for Municipalities in its four editions from 2017 to 2024 (see Table 1). The workshops were designed in such a way that the practical requirements for an indicator system were surveyed and ranked, supplemented with scientific standards, and, on this basis, each proposed indicator was discussed with regard to its further development. The practical requirements, as well as information on individual indicators that allowed conclusions to be drawn about the set of indicators and their application in general, gathered during this seven-year period of semi-structured group discussions, were interpreted in light of the research questions.

Table 1. Central events in the project SDG Indicators for Municipalities for collecting requirements and aspects for discussion from representatives of municipalities in Germany.

SDG Indicators for Municipalities Edition	Event	Date/Venue	Total Participants	No. of Municipalities	No. of Organisations
4th edition, 2024 Development phase: 2023–2024	Workshop	22nd of November 2024, virtual	64	57	3
	Workshop	11th of January 2024, virtual	84	61	4
3rd edition, 2022 Development phase: 2021–2022	Workshop	10th of November 2021, virtual	39	16	16
2nd edition, 2020 Development phase: 2019–2020	Workshop	1st of October 2019, Hanover	25	12	8
1st edition, 2018 Development phase: 2017–2018	Workshop	10th of October 2017, Hanover	27	13	5

The 239 participants, most of whom had attended several workshops, represented 84 unique municipalities. More than 60 percent of these municipalities are cities with more than 100,000 inhabitants, which means that the sample primarily reflects an urban perspective. The comments provided by participants touch on a wide range of issues including the following: the importance of specific aspects for indicator development ("We value that there are now more indicators for SDG 13 & 17 with data"), barriers that hinder the application of indicators ("The modal split survey is expensive and cannot be mapped at city district level"), and conceptual questions regarding the design of indicators or indicator groups ("Why is the number of staff working on climate protection not measured?"). Due to the application of Chatham House Rules during the workshops and the general objective of developing a set of indicators for all municipalities, it is not possible to retroactively attribute the feedback to individual municipalities or municipality types. The workshop minutes, notes, and related emails contained a total of 64 comments and suggestions that were suitable for the analysis conducted in this study.

The data analysis was carried out by systematically combining both sources, literature on indicator evaluation frameworks and primary data from group discussions. In a first step, the 45 criteria identified from seven IEFs were grouped into 16 criteria clusters (see Table 2). Those clusters are the result of a comparison of the IEFs (rows in Table 2), each containing similar criteria, albeit sometimes labelled differently and formulated with different nuances. While the "validity" criterion in the SDG Indicators for Municipalities, for example, primarily addresses internal validity, the SPICED framework ("Subjective, Participatory, Interpreted and communicable, Cross-checked and compared, Empowering, Diverse and disaggregated") adds that validity should be cross-checked by various stakeholders, thereby incorporating elements of an external validity check. These differences are mostly due to the different purposes and target groups of the individual IEFs and were taken into account accordingly when describing the final criteria (last column in Table 2). In a second step, each identified practical requirement from the workshops was assigned to a criterion cluster, e.g., data availability. Because a practitioner's statement could touch on more than one criterion, a second assignment was allowed. Both criteria and requirements were added up, resulting in a ranking. As soon as 10 criteria or requirements were identified from the IEFs and the practical feedback, those criteria/requirements were considered essential for future urban monitoring systems and were discussed along with the results of the literature review.

Table 2. Comparison of IEFs and practical requirements for identifying essential criteria for future urban monitoring. The IEFs are acronyms for individual evaluation criteria listed in each column. Due to sorting by sum of criteria, the acronyms are not in the correct order. The colour intensity indicates the frequency of mentions: the more frequently an aspect is addressed in the IEFs or in practice, the greener the sum and the more relevant the requirement is for future urban monitoring.

IEFs										
SDG Indicators for Municipalities (Bertelsmann Stiftung et al., 2024)	SMART* (Doran, 1981, modified by Selvik et al., 2021)	RACER (European Commission, 2023, first published in 2005)	CREAM (Schiavo-Campo & Tommasi, 1999)	SPICED (Roche, 1999)	ADAPT (Villanueva, 2011)	FABRIC (UK National Audit Office, 2001)	Description that summarises the key points of the respective IEF aspects	Sum of criteria from IEFs	Sum of requirements from municipalities	Requirements for future urban monitoring
Validity	Specificity	Relevant	Relevant	Cross-checked and compared	Thorough	Focused	The indicators are closely linked to the objectives/aims/goals to be reached. They are formulated precisely and the data allow consistent interpretation. The validity of assessment needs to be cross-checked and compared by different actors and methods.	7	26	Relevance and validity
Data availability	Time-based	Easy to monitor	Economic			Cost-effective	The indicator values are available at reasonable cost and burden in an appropriate time frame and coverage.	5	10	(Economic) data availability
Comprehensibility		Credible	Clear	Interpreted and communicable		Appropriate	The indicators and the validity of the indicators are credible for non-experts, unambiguous, and easy to interpret and communicate.	5	9	Comprehensibility
Control relevance	Relevancy		Adequate		Dynamic	Integrated	The indicators provide essential information for management, are integrated, and thus, allow improvement of the overall performance.	5	5	Impact-oriented management

Table 2. (Cont.) Comparison of IEFs and practical requirements for identifying essential criteria for future urban monitoring. The IEFs are acronyms for individual evaluation criteria listed in each column. Due to sorting by sum of criteria, the acronyms are not in the correct order. The colour intensity indicates the frequency of mentions: the more frequently an aspect is addressed in the IEFs or in practice, the greener the sum and the more relevant the requirement is for future urban monitoring.

IEFs									
SDG Indicators for Municipalities (Bertelsmann Stiftung et al., 2024)	SMART* (Doran, 1981, modified by Selvik et al., 2021)	RACER (European Commission, 2023, first published in 2005)	CREAM (Schiavo-Campo & Tommasi, 1999)	SPICED (Roche, 1999)	ADAPT (Villanueva, 2011)	FABRIC (UK National Audit Office, 2001)	Description that summarises the key points of the respective IEF aspects	Sum of criteria from IEFs	Sum of requirements from municipalities	Requirements for future urban monitoring
Data quality	Measurability	Robust against manipulation	Monitorable				The indicator data are measured accurately and reliably, enabling comparisons with other data.	4	2	
Function	Achievability				Active		The indicators can be differentiated in function and provide adequate information on actions, impacts, and states.	3	3	
Participation				Participatory	Participatory		The indicators are developed by and with those best placed to assess them or affected by the interventions.	3		
Acceptance		Accepted					The indicators are widely accepted, and their role and responsibilities are well defined.	2	5	
Completeness						Balanced	The indicators cover all significant areas of work.	2	4	
Stability						Robust	The indicators are able to withstand organisational changes and remain stable.	2	2	

Table 2. (Cont.) Comparison of IEFs and practical requirements for identifying essential criteria for future urban monitoring. The IEFs are acronyms for individual evaluation criteria listed in each column. Due to sorting by sum of criteria, the acronyms are not in the correct order. The colour intensity indicates the frequency of mentions: the more frequently an aspect is addressed in the IEFs or in practice, the greener the sum and the more relevant the requirement is for future urban monitoring.

IEFs										
SDG Indicators for Municipalities (Bertelsmann Stiftung et al., 2024)	SMART* (Doran, 1981, modified by Selvik et al., 2021)	RACER (European Commission, 2023, first published in 2005)	CREAM (Schiavo-Campo & Tommasi, 1999)	SPICED (Roche, 1999)	ADAPT (Villanueva, 2011)	FABRIC (UK National Audit Office, 2001)	Description that summarises the key points of the respective IEF aspects	Sum of criteria from IEFs	Sum of requirements from municipalities	Requirements for future urban monitoring
Current relevance					Adaptive		The indicators reflect changes and correspond to the current state of science and practice.	2	2	
				Diverse and disaggregated			The indicators are diverse and disaggregated from a range of groups.	1	14	Diversity and aggregation level
Compatibility							The indicators are compatible with indicators from other sets.	1	11	Compatibility
Manageability							The indicators are not redundant. The number of indicators is therefore manageable.	1	9	Manageability
				Empowering			The process of setting and assessing indicators is empowering in itself and allows groups and individuals to reflect critically on their changing situation.	1	3	
				Subjective			The indicators are based on unique insights of informants.	1	1	

Notes: * SMART is presented as one specific set of criteria, however, many versions of the SMART acronym exist, where the letters can refer to other aspects or criteria; the analysis presented in this table is based on the modification by Selvik et al. (2021), which is designed to assess the quality of indicators; in his original publication, Doran (1981) was specifying that objectives should be specific, measurable, assignable, realistic and time-related.

3. Results and Discussion

Following the methodological approach, the literature review identified seven relevant IEFs, designed and used to guide the measurement of SUD. Due to their different origins and application purposes, each IEF has distinct characteristics, resulting in various strengths and weaknesses for indicator evaluation and selection in the context of SUD. Regardless of whether an IEF is more scientific or policy-oriented, it must be emphasised that IEFs are becoming more important as the number of available indicators and data grows, while administrative capacity in local and regional governments remains constant or even declines. In general, by systematically evaluating and selecting SUD indicators using clear criteria, cities can more effectively measure what matters, make and legitimise informed decisions, and drive effective change toward sustainability targets (e.g., Lynch et al., 2011). Choosing the right indicators remains a crucial step in the monitoring process, as the entire purpose of introducing an monitoring and evaluation system, which is directly linked to the increasing number of municipal sustainability strategies (Gieseler et al., 2024), is undermined if the indicators are not selected appropriately.

The SDG Indicators for Municipalities have evolved on the basis of the research and discussions that are also the subject of this article, which is why they address many criteria but have their own focus as a result of deliberation processes with various stakeholders (Bertelsmann Stiftung et al., 2024). The most prominent IEF is probably SMART ("Specificity, Measurability, Achievability, Relevancy, Time-based"), although many versions of the SMART acronym exist, where the letters sometimes refer to different aspects or criteria (Selvik et al., 2021). The scheme was originally developed for the purpose of guiding the development of goals and targets (Doran, 1981), which is why the criteria may not always be appropriate for indicator development. Furthermore, the practical application of the SMART principles is still problematic, as the criteria are interpreted differently (Climate-Eval Community of Practice, 2015). Nevertheless, being one of the first assessment tools (Doran, 1981), the SMART criteria have been widely applied by scholars and practitioners (e.g., Lynch et al., 2011; Namavar et al., 2023; Simon et al., 2016). Similar to the SMART criteria, but with a clear focus on indicator evaluation, the RACER framework is becoming increasingly established—not least because it has been developed, applied, and recommended by the European Commission. The RACER criteria originate from the European Commission's Impact Assessment Guidelines and are used in the EU to evaluate the quality of indicators, particularly in the context of the circular economy transition and other policy initiatives, such as the monitoring of natural resource use in Germany (Gerdes et al., 2011; Papageorgiou et al., 2025). The RACER framework is also increasingly finding its way into monitoring and evaluation frameworks for SUD (e.g., Angelidou & Politis, 2024). With its focus on easy monitoring and acceptance of indicators, the IEF is particularly interesting for practical users. CREAM principles ("Clear, Relevant, Economic, Adequate, Monitorable") were originally designed to select good performance indicators (Schiavo-Campo & Tommasi, 1999). However, its generally softer criteria make it possible to assess quantitative and qualitative indicators at the same time (Zall Kusek & Rist, 2004). These broader guidelines, expressed through aspects such as "adequate" and "monitorable" rather than "specific" and "measurable" as in the SMART scheme, can be beneficial when the quantification of processes and changes is difficult, but qualitative approaches allow progress to be assessed. Therefore, CREAM is recommended for complex contexts, such as the measurement of water resources management or gender-sensitive evaluations (e.g., Bertule et al., 2017; OECD, 2021). In contrast to all other IEFs, the SPICED principles focus on how indicators should be used rather than on how they should be developed—therefore, it is recommended to combine this framework with others (Climate-Eval Community

of Practice, 2015). These principles emphasise indicators that empower change (Roche, 1999). Originally developed for the impact assessment of development programmes, the SPICED framework is also used for assessing cultural ecosystem service indicators, for example (Hernández-Morcillo et al., 2013). With the ADAPT IEF ("Adaptive, Dynamic, Active, Participatory, Thorough"), Villanueva (2011) proposes a more flexible, alternative approach to SMART indicators that is more conducive to climate change adaptation and climate risk management and is consequently being used increasingly in this area. According to this IEF, indicators should measure actions rather than states. The criteria of the FABRIC IEF ("Focused, Appropriate, Balanced, Robust, Integrated, Cost-effective") set the property of performance information that would be derived from the selected indicators (UK National Audit Office, 2001). Used primarily in institutional environments, this framework focuses on indicator assessments that allow the integration of management processes and the improvement of efficiency. Since all IEFs, albeit with varying attributes, refer extensively to indicator evaluation and are applied accordingly in SUD and its sub-fields, they are used to derive future requirements for indicators. The fact that almost all criteria are also addressed in the feedback from the municipalities further validates the selection of IEFs included in this analysis.

The 16 clusters that were created as part of this study contain only five criteria that are unique to a single IEF. This suggests the existence of universal quality criteria for indicator catalogues; these criteria can always be applied, regardless of the thematic focus or target group. Three of these five criteria are frequently discussed by practitioners and are classified as particularly important. On closer inspection, however, this is hardly surprising, as both the SDG Indicators for Municipalities and the SPICED framework consider not only how indicators should be developed but also how they should be applied, a topic which is highly relevant to municipalities. There may be various reasons why practitioners do not mention participatory indicator development, such as the fact that participation may already be perceived as part of a broader strategic process in the municipality, or the fact that responsibility for the selection of indicators may rest with administrative experts, while responsibility for the interpretation of the data is shared by all.

The overall comparison identified seven central criteria required by theory and practice: Relevance and validity, comprehensibility, impact-oriented management, and economic data availability are important criteria in both domains. In addition to technical aspects, such as data quality, function, or interdependencies of indicators, the designers of future monitoring systems should also provide the flexibility ("diverse and disaggregated"), compatibility, and manageability demanded by practitioners. Since all criteria are interlinked in many ways, elements from other frameworks are also considered in the following individual descriptions and discussions.

3.1. Relevance and Validity

There is no standardised definition or framework for SUD. In fact, the 2030 Agenda can be seen as the first framework designed to monitor progress in a systematic way. As a globally negotiated political consensus, the 2030 Agenda serves as a universal roadmap, but it poses challenges for local governments: Some of the targets are abstract, often unquantified, and primarily aimed at the national level in all countries equally (Valencia et al., 2019). For local authorities in individual countries, the development policy objectives and the means of implementation must be translated or, in some cases, omitted (López Chao et al., 2020; Nagy et al., 2018). In addition, the availability of small-scale data must be reviewed. Therefore, the SDG Indicators for Municipalities project included a comprehensive relevance check (Bertelsmann Stiftung et al., 2024). For this purpose, the 169 targets of the 17 SDGs were examined to determine whether they needed to be broken

down into sub-targets in order to enable a consistent assessment of their relevance for German municipalities. This was necessary for some targets, as components with different content may need to be assessed differently during the problem check or the task check. These checks answer the following questions: Does the target or sub-target describe a problem occurring in German municipalities, and is the local level capable of taking appropriate action? The final step is to consider whether the municipality can also generate impact in the Global South, for example through partnerships. Unfortunately, SDG targets pose a fundamental problem, as their wording leaves room for interpretation, making the results of the relevance check dependent on the specific interpretation chosen. For example, it is initially unclear what is meant by the "modern energy services" referred to in SDG 7.1: In some countries, nuclear power plants could also be considered modern energy sources; however, this would contradict the basic principle of "planet," which explicitly calls for the protection of the planet from harm and action against climate change. In order to classify SDG 7.1 as a relevant sub-goal, "modern" was equated with clean and renewable. In contrast, the desertification mentioned in SDG 15.3.1 was not "translated" as soil degradation. As a result, this sub-target is not relevant for German municipalities, but is nevertheless relevant for municipalities in the Global South.

Despite localisation efforts, the SDG Indicators for Municipalities inherit the limitations of the orientation framework and therefore raise practical challenges. Lack of clarity with regard to the scope of the goal, different or missing time horizons, conflicts between ambition levels at global, national, and regional levels, etc.—the list of issues is long. According to feedback from municipalities, indicators often either fall short of the target, as seen with the recycling rate for SDG 12.5 ("By 2030, substantially reduce waste generation through prevention, reduction, recycling and reuse"), or exceed it, as with the premature mortality indicator for SDG 3.4 ("By 2030, reduce by one third premature mortality from non-communicable diseases through prevention and treatment and promote mental health and well-being"). For municipalities, the question often arises as to how an issue should be measured in order to do justice to sustainability. Municipal services can be seen as socially sustainable per se, but questions remain regarding how sustainable the services are in themselves (e.g., sustainable construction of a daycare centre) and what access the population has to the city's services (e.g., physical access, proximity, affordability, safety, etc.). Moreover, certain subject areas in the 2030 Agenda that are underrepresented in the municipal context include, importantly, culture, sport, digitalisation, and land management.

A post-2030 agenda must contain clear and measurable targets, using indicators that are widely accepted and easy to implement. For example, SDG 13 could be substantiated with net-zero carbon dioxide emissions by mid-century and operationalised climate resilience targets (Fuso Nerini et al., 2024). Many of the issues raised above could be addressed by applying a systematic framework such as DPSIR ("Drivers, Pressures, State, Impact, and Response") to a post-2030 agenda in order to promote a better understanding of the interrelations within and between goals, targets, measures, and impacts (Carr et al., 2007). DPSIR can be useful for assessing sustainable development at global, national, and local levels, but this requires finding a way to incorporate the aggregate effects of local and also informal responses to drivers, pressures, and conditions. This involves linking qualitative and quantitative data, considering objective and subjective states, and assessing target synergies and conflicts (Jossin & Peters, 2022). Rethinking and embracing the complexity and disorder of the connection between human well-being and the environment is an enormous challenge that cannot be fully resolved in a new framework for sustainable development. However, it becomes clear that scientific evidence and systems logic will need to be taken into greater account in any

forthcoming policy consensus. Implementing sustainable development, especially in the municipalities, is difficult in itself—an ambiguous framework does not contribute to greater acceptance and acceleration.

3.2. Compatibility

Since there is no uniform framework, but rather many guidelines and regulations, local authorities must ensure vertical and horizontal integration in order to ensure accountability. Vertically, municipalities are asking which policy levels they can use for comparison with their own values (e.g., county or national average). Horizontally, they are considering which other SUD frameworks, policies, and planning instruments their indicators can be compared with. In this context, inter-municipal comparability must also be mentioned, as it enables best practices to be learned on the basis of indicators. To address these issues, the SDG Indicators for Municipalities project emphasises transparency with regard to indicator origin and provides compatibility checks with indicators from other frameworks. Compatibility is checked and disclosed in two ways: firstly, based on the consistency of the indicators in terms of content, and secondly, based on definitional conformity. If two indicators share the same definition or calculation, the indicator is considered fully vertically integrated, meaning that aggregation or disaggregation should be possible. While the employment rate (SDG 8.5) and nature conservation areas (SDG 15.5) are defined largely consistently throughout various entities and can therefore be easily compared, different parameters usually apply to the cycle path network (SDG 11.2) and to the definition of homelessness (SDG 1.3), making it difficult to compare the data.

From a practical perspective, municipalities are often faced with the challenge of deciding which framework to use as a guideline when setting up a municipal sustainability management system. The 2030 Agenda is chosen not least because it is considered to be vertically integrated and connectable. However, some cities also use other orientation frameworks such as the "Economy for the Common Good" (Felber, 2019), and seek interfaces to other frameworks, such as the SDGs. Municipalities are faced above all with the questions of how the indicators contribute to the requirements of current or future legal frameworks and how they can ensure the comparability of indicators that are already being used in the municipality, e.g., in sectoral strategies. For indicators already in use, there are often minimal definitional deviations that show a different status quo or a different development. This also applies at the data level: In many municipalities, their own data deviate from the data contained in higher-level official statistics. If the definitions of indicators match, variances with regard to the data source can be identified more easily.

In general, compatibility and standardisation of indicators counteract the principle of necessity to adapt indicators to the specific needs of municipalities (Michalina et al., 2021) and data availability. This can be a major obstacle to effective impact measurement and the integration of indicators into individual management processes. In addition, compatibility is a difficult criterion for the selection of indicators if these are developed in a participatory manner—e.g., together with citizens. Hansson et al. (2019) therefore suggest decentring global measurability and comparability, as pre-defined indicator sets inhibit creativity when formulating indicators that can promote positive behaviour or behavioural change. However, standardised indicator language provides transparency, reliability, harmonisation of indicators, and comparability of results (Ruan & Yan, 2022). Furthermore, there is a risk that local interpretations can lead to unambitious measures where the SDGs are primarily used for branding, rather than as guidance for transformational change (Valencia et al., 2019). Expert-led and citizen-led approaches—the two methodological paradigms

that have evolved—and their respective characteristics are widely discussed in the literature (Turcu, 2013). For future monitoring systems, this implies an integration of compatible and adapted indicators whose origin and comparability are presented clearly and transparently.

3.3. Manageability

Although a catalogue of indicators that fully reflects the content of all 169 targets should be the ultimate goal, such a catalogue would probably have to consist of a multitude of indicators and would therefore go beyond the scope of municipal sustainability monitoring. Local authorities should be able to use the catalogue to take stock of their sustainable development as comprehensively as possible, but this process needs to be manageable and realistically feasible. For this reason, the focus of the SDG Indicators for Municipalities is on those targets or sub-targets that address key problems and challenges in German municipalities or in municipalities in the Global South and that can be solved or overcome by German municipalities—at least partially and, above all, measurably—through their own expertise. Generally, the indicator catalogue should contain no more indicators than there are relevant targets and sub-targets, assuming that some targets cannot yet be methodically measured through indicators and that some targets require different indicators to adequately reflect them. In SDG 3.8 ("Achieve universal health coverage, including financial risk protection, access to quality essential health-care services and access to safe, effective, quality and affordable essential medicines and vaccines for all"), for instance, various medical facilities would need to be covered to gain an impression of the local health-care situation. However, measuring the municipal contribution to providing access to affordable essential medicines and vaccines for all is challenging, which is why no adequate indicators are expected at this point. Indicators that aggregate impacts or cover a range of measures, such as "nitrate in groundwater" or "operating sites with environmental and sustainability certificates," are in turn linked to so many targets that they cannot be used to measure individual targets.

Many municipalities find large catalogues of indicators difficult to manage and communicate ("We have considerable interest in simple and clear solutions with an even distribution of a small number of indicators per target"). In practice, municipalities often deviate from the SDG logic and focus on selected strategic goals in their own fields of action with a single key indicator per area. However, this approach conflicts with the principle of completeness, as important aspects or target relationships may be neglected. When structuring a monitoring report based on the SDGs, this is particularly evident in SDG 11, the most important SDG for SUD, as it covers several areas such as housing, transport, cultural heritage, and natural disasters, which cannot be adequately represented by a single indicator.

In general, the length and depth of the indicator catalogue probably represents one of the most difficult balancing acts between scientific and practical requirements (Verma & Raghubanshi, 2018). From a theoretical perspective, the indicator set should strive for completeness in order to cover all relevant aspects of SUD and to be able to identify adequately interactions between indicator trends (Miller et al., 2013). However, instead of comprehensive metrics of complex urban sustainability issues, policy-makers in particular need to see indicators as supporting tools for informed decision-making (Valencia et al., 2019). Furthermore, a high number of indicators poses administrative challenges related to data collection. Taking regular updates of the data into account, data collection becomes a significant economic factor (Chen et al., 2025). Therefore, a smaller number of multi-purpose indicators that contribute to various targets could help to keep a future

indicator catalogue manageable. The same applies to composite indices, but their complex calculation and weighting may hinder validity and comprehensibility (Rodrigues & Franco, 2020; Tanguay et al., 2010).

3.4. Comprehensibility

As the level of government closest to citizens, municipalities need comprehensible indicators, as opposed to higher-level catalogues focused on measures of control. SUD monitoring must be accepted by a wide range of local stakeholders. However, data interpretation, as well as governance characteristics and processes, can often only be understood by experts (e.g., municipal roles and responsibilities). Comprehensibility is a cross-cutting criterion and can be evaluated in relation to all other criteria, across various purposes and target groups. The SDG Indicators for Municipalities evaluate the following statement: "An indicator is considered 'plausible' if both the statement of the indicator itself and the reference to the respective target or sub-target are comprehensible." Comprehensibility may be impaired if an indicator is based—usually for reasons of data availability—on a definition or calculation that is not unequivocally clear, if it is associated with scientific and technical model assumptions, or if it can generally only be understood with (legal) background knowledge. As a result, it is not always possible to identify a clear link to the target at first glance. Accordingly, indicators are assessed differently. The "rental prices" indicator can be easily understood, and the approach to measuring access to affordable housing (SDG 11.1) is clear. In contrast, the indicator "landscape quality," a composite index that measures a wide range of factors related to the cultural impact of humans on ecosystems on a 7-level scale (technical term: hemeroby), and therefore contributes to reducing the degradation of natural habitats (SDG 15.5), is less easy to comprehend. However, as one of the few proxies for biodiversity loss, the hemeroby index is a very valuable indicator. It is representative of the phenomenon that whenever official statistics are exhausted and scientific measurement concepts (must) be used, comprehensibility may be impaired.

In practice, in addition to the comprehensibility of the indicators and their validity, the interpretation of data is much discussed. While for some indicators an increase or a decrease is clearly positive in terms of sustainable development, this one-dimensional assessment is usually not possible for all indicators. The indicator "SGB II and XII rate" (SDG 1.3), for example, describes the proportion of eligible social welfare recipients in German municipalities according to the Social Insurance Code (SGB) II and XII. Both a decrease and an increase of values could be considered positive in terms of sustainable development: A falling rate could indicate that fewer people are dependent on social assistance, whereas a rising rate could mean that more people are receiving help. There are several other examples of sustainability indicators, especially in the social dimension, that can be interpreted ambiguously.

For a post-2030 agenda, the criterion of comprehensibility is of particular importance, as it bridges the gap between science, planning practice, administration, and politics. Thus, the complexity of indicators must be reduced and standardisation operations applied (Schetke et al., 2012). However, the more comprehensible an indicator and its characteristics needs to be, the more likely that less innovative measurement concepts can be used to close thematic gaps and that fewer administrative specifics can be taken into account, which may hamper impact assessments. Furthermore, the ability to measure sustainability dimensions in depth may be affected (Verma & Raghubanshi, 2018). Therefore, the usage and publication of indicator catalogues should be accompanied by comprehensive communication concepts and guidelines or training to ensure that everyone can follow and interpret the indicators and underlying data. In general, the applicability and acceptability of

indicator catalogues as well as the political influence on final indicator decisions will always be an issue (Mori & Christodoulou, 2012). Against this background, transparent methodologies in future measurement concepts and sustainability reports are essential.

3.5. Impact-Oriented Management

Monitoring is primarily intended to enable the continuous, systematic collection and analysis of data, with the goal of measuring progress toward achieving specific goals and targets. Embedded in a municipal sustainability management system with strategies and implementation plans, the data can be contextualised and analysed before and after an intervention in order to measure the potential or actual impact. The SDG Indicators for Municipalities aim to provide only indicators that enable impact-oriented management. However, two observations were made during the further development of the indicators, which were confirmed in the group discussions: Firstly, some indicators allow better derivation of necessary measures than others. Secondly, some indicators already imply a normative target achievement path that is interpreted into the target for municipalities. Both observations are based primarily on the underlying impact logic. Indicators that describe inputs and outputs and are therefore closer to the administrative realities can be more easily translated into actions. Outcome and impact indicators, on the other hand, require an assessment of the interventions that have led to certain effects. For example, the development of the impact indicator "ratio of employment rates of women and men" (SDG 5.1 "End all forms of discrimination against all women and girls everywhere") depends on many factors in the individual and overall societal environment. The "proportion of women on the city council, municipal council or district council," on the other hand, is a simple lever to "Ensure women's full and effective participation and equal opportunities for leadership at all levels of decision-making in political, economic and public life" (SDG 5.5).

In practice, many municipalities naturally prefer the latter indicators with clear links between targets, measures, and indicators. Another example is the measurement of human resources, which often poses a particular challenge in German municipalities due to a shortage of staff ("Why is the number of staff working on climate protection not measured? In the absence of dedicated sustainability staff, these staff often have to keep an eye on these issues and are therefore very relevant."). However, such indicators often only assume causal relationships without being able to quantify the exact impact, and this limits the problem-solving capacity of municipalities.

The SDGs and their indicators do not primarily address the quality of policies and plans nor the quality and effectiveness of their implementation (Valencia et al., 2019). Usually, the indicators represent either the beginning or the end of a cause–effect relationship and therefore rarely allow conclusions to be drawn about the impact of a particular measure. Despite experimenting with calculations and indexations that are combined with context, correlation, and sensitivity analysis, differing and inconsistent results can lead to misconceptions about the ability to assess SDG impacts (Janoušková et al., 2018). The development of valid sustainability impact assessments requires a comprehensive list of indicators and a range of weightages that vary according to context (Kaur & Garg, 2019). Synthesising completed impact assessments into standardised impact pathways that can be applied across different municipalities and situations has the potential to enhance monitoring systems in a post-2030 agenda.

3.6. Diversity and Aggregation Level

Indicators can be designed and analysed across multiple dimensions, including gender, age, education, sector, and administrative unit. Some of these also depend on the local operational level. Targeted assessments are useful as the challenges and policy impacts differ between groups. This applies in particular if the indicators reveal trends that require different approaches. Under the central, transformative promise of the 2030 Agenda, "Leave no one behind," many targets and indicators within the SDGs predefine a distinctive consideration of groups. In the SDG Indicators for Municipalities, this is the case for poverty rates (SDG 1.3) that are differentiated for children, adolescents, and the elderly. Employment rates (SDG 8.5) are broken down by age and origin, energy balances of buildings (SDG 11.b) by new and existing buildings, water and energy consumption (SDG 12.2) by private households and industry, and greenhouse gas balances (SDG 13.2) by multiple sectors.

This differentiation is certainly supported by local authorities, but there is little consensus on the question of which subject areas require more attention. In many cases, municipalities want to sub-categorise indicators according to their administrative structure, and other municipalities because of their direct influence. Although the effect of policy instruments and sensitisation efforts for all groups outside the core administration in fact often remains unclear, the demarcation is still very difficult. For example, how should municipal companies and shareholdings such as municipal utilities, waste management, and local transport companies be addressed, given that there are at least indirect opportunities to exert influence? In the latter example, there are also strong differences between the sub-national aggregation levels. Depending on the respective state constitution, cities, counties, and towns are often responsible for different remits, which is why a so-called "scale blindness" is noticeable in the indicators—i.e., not all indicators are equally relevant for all types of local authorities. Furthermore, administrative borders may distort the measurement object and/or not take spill-over effects into account.

Generally, indicators should be diversifiable to address the most vulnerable groups and scalable to the governance level with maximum influence. The more tailored an indicator is, the more likely the measurement and results will be accepted for supporting implementation (Frare et al., 2020; Hansson et al., 2019). At the same time, it can be argued that indicators should not supersede other local measures of change, but complement and reinforce them. By definition, indicators point to facts that are not directly measurable or to broader issues. They must therefore not be used as definitive measures of specific interventions, but must be understood as guidelines for assessing progress (Klopp & Petretta, 2017). Since the diversity of indicators is also often the result of data unavailability (Ruan & Yan, 2022), and individual adjustments of indicator scope counteract the compatibility principle, a post-2030 agenda should pursue a flexible monitoring system with core and municipality-specific indicators.

3.7. (Economic) Data Availability

Given their limited resources, local authorities in particular are dependent on data availability and on cost-effective data collection for SUD. Knowing about the data availability bottleneck, the ambition of the SDG Indicators for Municipalities is to provide comprehensive data for all indicators for German municipalities with more than 5,000 inhabitants. In contrast to many other indicator sets, the project thus pursues a top-down approach. This is because most data is not based on voluntary reports from the

municipalities, but on official statistics from the federal or national government or overarching research projects. However, some indicators still require decentralised collection by the municipality. Indicator data relying on the collection based on this bottom-up approach, i.e., voluntary reporting by municipalities, could often not be published in the past, as the reports were either inadvertently or deliberately incomplete—usually for political reasons. This was the case, for example, for indicators on greenhouse gas balances (SDG 13.2) and city-to-city partnerships in the Global South (SDG 17.16).

Although municipalities value the provision of data on SUD, there are two major obstacles in practice: First, data vary between different statistical authorities, especially data on population, and this often leads to confusion. Second, data availability at the local level is still generally poor in temporal and spatial resolution (Hák et al., 2016). Most of the data are only available at the level of large cities and districts, have a time lag of up to two years, and are not collected annually. Accordingly, this is criticised when important indicators such as basic medical care close to home (SDG 3.8) cannot be provided for small, rural municipalities or when data on the expansion of renewable energies (SDG 7.2) are not up-to-date. Furthermore, the appropriate scale is discussed, especially for indicators that should also reflect the situation at the neighbourhood level (e.g., noise pollution) or in surrounding areas of the municipality (e.g., recreational areas). In contrast, when municipalities have to collect their own data, individual assessments are required, and some of them can only be operationalised with considerable financial and human resources—assuming detailed measurement concepts are available. Together with the time and economic pressure that data collection puts on local administrations, the data bottleneck can lead to indicators not being reported at all, impact indicators being replaced by input indicators (e.g., expenditure on specific measures), or a certain standard of data quality not being met.

For an economic, disaggregated data availability, advancements in digitalisation, artificial intelligence, and remote sensing, as well as strong capacity building, should be taken into account for post-2030 agenda monitoring systems (Lehner et al., 2018). In this context, smart and sustainable urban development should be considered jointly (Rathore et al., 2016), and local advocacy must be promoted, e.g., by linking indicator and budget development, which often attracts attention (Klopp & Petretta, 2017).

4. Limitations of the Study

This study tries to obtain a general overview of key aspects to consider in future indicator systems by means of a literature review of IEFs and related studies, complemented by insights from group discussions with municipalities located in Germany. Both sources have their limitations: The IEFs were developed for different purposes, hence they may not take the specifics of SUD governance frameworks, the planning system, or the local situation into account. Similarly, the discussion formats with the municipalities were originally not designed to systematically identify requirements for future urban monitoring approaches. Consequently, all comments and suggestions have been carefully checked for indications in this direction. However, it cannot be ruled out that some remarks have been misinterpreted. The amount of criteria and the remarks from both sources, together with the comprehensive literature review, are intended to eliminate major biases in the discussion. Enhanced by methodological approaches that analyse the concrete application of the various IEF criteria on site, e.g., the analysis of indicator sets in Voluntary Local Reviews, future research could further refine the results of the present study, potentially leading to a proposal for a novel IEF.

5. Conclusion

Sustainable urban development combines two complex disciplines that cover the entire spectrum of spatial impacts from socially just, ecologically sound, and prosperous governance. Although not primarily designed for this field or this policy level, the 2030 Agenda is seen as a framework that can guide municipalities as they strive to become inclusive, safe, resilient, and sustainable. Future monitoring systems must overcome the inherent limitations of the SDGs for SUD while meeting central theoretical and practical requirements. This article sheds light on experiences from SUD monitoring approaches and the SDG Indicators for Municipalities, aiming to provide valuable conclusions for measuring the impacts of municipal sustainability management.

Seven key requirements were identified, highlighting the need for balance in the selection of indicators depending on the perspectives and priorities, whether for a post-2030 agenda, a conceptual indicator catalogue such as the SDG Indicators for Municipalities, or local government frameworks. Relevance and validity, economic data availability, comprehensibility, and impact-oriented management are important criteria for theory and practice. In addition to technical aspects such as data quality, function, or interdependencies of indicators, the designers of future monitoring systems should also provide the flexibility, compatibility, and manageability demanded by practitioners. These criteria are not without contradictions. Indicators for SUD must reflect the environmental, social, and economic dimensions of sustainability while taking into account different conceptions of resilience, inclusion, and governance. When evaluating and selecting indicators, local and regional governments aim to identify the most suitable measuring variables that explain and guide developments for a wide range of stakeholders. In this respect, any remaining ambiguities in the identified and discussed criteria can serve as the flexible elements needed locally to facilitate policy-making for SUD. Depending on the governance structures and the local planning and participation culture, this can be achieved in a bottom-up or top-down manner, with a more scientific or practical orientation, and in a customised or interconnected way. A post-2030 agenda must reflect this flexibility while ensuring a universal and valid language for indicators and an easy derivation of actions. From a technical perspective, the many trade-offs between the criteria can be overcome with smart monitoring approaches: A set of core and municipality-specific indicators could ensure compatibility and adaptability; digital collection and capacity building could enhance data availability; communication approaches could promote comprehensibility; and multi-purpose indicators could limit the unmanageable length of an indicator catalogue. Above all, however, the further development of a clear, unambiguous, and relevant orientation framework that incorporates a strategic and operational nexus of targets, measures, and indicators remains of utmost importance.

Acknowledgments

The authors thank the Bertelsmann Stiftung for funding the development and enhancements of the SDG Indicators for Municipalities and therefore providing the basis for the consideration of an improved monitoring approach. We would also like to thank three anonymous reviewers and the editors of the thematic issue for their helpful comments.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

Further information and data on the SDG Indicators for Municipalities are available at https://sdg-portal.de/en

References

Angelidou, M., & Politis, C. (2024). An integrated framework to monitor progress of sustainable urban development strategies, addressing societal, ecological and innovation dilemmas. *Societal Impacts*, *4*, Article 100075. https://doi.org/10.1016/j.socimp.2024.100075

Benedek, J., Ivan, K., Török, I., Temerdek, A., & Holobâcă, I. H. (2021). Indicator-based assessment of local and regional progress toward the Sustainable Development Goals (SDGs): An integrated approach from Romania. *Sustainable Development*, *29*(5), 860–875. https://doi.org/10.1002/sd.2180

Bertelsmann Stiftung, Bundesinstitut für Bau-, Stadt- und Raumforschung, Deutscher Landkreistag, Deutscher Städtetag, Deutscher Städte- und Gemeindebund, Deutsches Institut für Urbanistik, Engagement Global, ICLEI European Secretariat, Kommunale Gemeinschaftsstelle für Verwaltungsmanagement, Landesarbeitsgemeinschaft Agenda 21 NRW, Rat der Gemeinden und Regionen Europas/Deutsche Sektion, & Rat für Nachhaltige Entwicklung. (2024). *SDG-Indikatoren für Kommunen: Indikatoren zur Abbildung der Sustainable Development Goals der Vereinten Nationen in deutschen Kommunen* (4th ed.). https://doi.org/10.11586/2025065

Bertule, M., Bjørnsen, P. K., Costanzo, S. D., Escurra, J., Freeman, S., Gallagher, L., Kelsey, R. H., & Vollmer, D. (2017). *Using indicators for improved water resources management: Guide for basin managers and practitioners*. UN Environment – DHI Centre on Water and Environment; World Wildlife Fund (WWF-US); Conservation International (CI); Luc Hoffmann Institute (LHI); University of Maryland Center for Environmental Science (UMCES); Global Environment Facility (GEF).

Carr, E. R., Wingard, P. M., Yorty, S. C., Thompson, M. C., Jensen, N. K., & Roberson, J. (2007). Applying DPSIR to sustainable development. *International Journal of Sustainable Development & World Ecology*, *14*(6), 543–555. https://doi.org/10.1080/13504500709469753

Chen, L., Shuai, C., Chen, X., Sun, J., & Zhao, B. (2025). Toward sustainable urban development: Identifying principal SDG indicators for Chinese cities. *Advanced Sustainable Systems*, *9*(4), Article 2400894. https://doi.org/10.1002/adsu.202400894

Climate-Eval Community of Practice. (2015). *Good practice study on principles for indicator development, selection, and use in climate change adaptation monitoring and evaluation*. https://www.adaptation-fund.org/wp-content/uploads/2017/08/DB_CCA-MandE-Indicators-Good-Practice-Study.pdf

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review*, *70*(11), 35–36.

European Commission. (2005). *Impact assessment guidelines*.

European Commission. (2023). *'Better regulation' toolbox 2023*.

Felber, C. (2019). *Change everything: Creating an economy for the common good*. Zed Books.

Frare, M. B., Clauberg, A. P. C., Sehnem, S., Campos, L. M. S., & Spuldaro, J. (2020). Toward a sustainable development indicators system for small municipalities. *Sustainable Development*, *28*(5), 1148–1167. https://doi.org/10.1002/sd.2065

Fuso Nerini, F., Mazzucato, M., Rockström, J., van Asselt, H., Hall, J. W., Matos, S., Persson, Å., Sovacool, B., Vinuesa, R., & Sachs, J. (2024). Extending the Sustainable Development Goals to 2050: A road map. *Nature*, *630*, 555–558. https://doi.org/10.1038/d41586-024-01754-6

Gebara, C. H., Thammaraksa, C., Hauschild, M., & Laurent, A. (2024). Selecting indicators for measuring progress towards sustainable development goals at the global, national and corporate levels. *Sustainable Production and Consumption*, *44*, 151–165. https://doi.org/10.1016/j.spc.2023.12.004

Gerdes, H., Bassi, S., Portale, E., Mazza, L., Srebotnjak, T., & Porsch, L. (2011). *D2.2 final report: Evaluation of indicators for EU policy objectives*. Institute for European Environmental Policy; Fondazione Eni Enrico Mattei; Ecologic Institute. https://www.ecologic.eu/14095

Gieseler, H., Ruddek, A., & Peters, O. (2024). Kommunales Nachhaltigkeitsmanagement in der Praxis—Eine Synthese von Potenzialen und Herausforderungen aus lokalen Nachhaltigkeitsanalysen. In K. Butzer-Strothmann & F. Ahlers (Eds.), *Kommunales Nachhaltigkeitsmanagement: Ein integrativer Ansatz mit Fokus Wirtschaft am Beispiel der Stadt Hannover* (pp. 355–370). Springer. https://doi.org/10.1007/978-3-662-67916-6_22

Giles-Corti, B., Lowe, M., & Arundel, J. (2020). Achieving the SDGs: Evaluating indicators to be used to benchmark and monitor progress towards creating healthy and sustainable cities. *Health Policy*, *124*(6), 581–590. https://doi.org/10.1016/j.healthpol.2019.03.001

Hák, T., Janoušková, S., & Moldan, B. (2016). Sustainable Development Goals: A need for relevant indicators. *Ecological Indicators*, *60*, 565–573. https://doi.org/10.1016/j.ecolind.2015.08.003

Hansson, S., Arfvidsson, H., & Simon, D. (2019). Governance for sustainable urban development: The double function of SDG indicators. *Area Development and Policy*, *4*(3), 217–235. https://doi.org/10.1080/23792949.2019.1585192

Hernández-Morcillo, M., Plieninger, T., & Bieling, C. (2013). An empirical review of cultural ecosystem service indicators. *Ecological Indicators*, *29*, 434–444. https://doi.org/10.1016/j.ecolind.2013.01.013

Janoušková, S., Hák, T., & Moldan, B. (2018). Global SDGs assessments: Helping or confusing indicators? *Sustainability*, *10*(5), Article 1540. https://doi.org/10.3390/su10051540

Jossin, J., & Peters, O. (2022). Sustainable Development Goals (SDG) indicators for municipalities: A comprehensive monitoring approach from Germany. *Journal of Urban Ecology*, *8*(1), Article juac020. https://doi.org/10.1093/jue/juac020

Kaur, H., & Garg, P. (2019). Urban sustainability assessment tools: A review. *Journal of Cleaner Production*, *210*, 146–158. https://doi.org/10.1016/j.jclepro.2018.11.009

Klopp, J. M., & Petretta, D. L. (2017). The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. *Cities*, *63*, 92–97. https://doi.org/10.1016/j.cities.2016.12.019

Koch, F., Beyer, S., & Chen, C. Y. (2023). Monitoring the Sustainable Development Goals in cities: Potentials and pitfalls of using smart city data. *GAIA – Ecological Perspectives for Science and Society*, *32*(Suppl. 1), 47–53. https://doi.org/10.14512/gaia.32.S1.8

Lehner, A., Erlacher, C., Schlögl, M., Wegerer, J., Blaschke, T., & Steinnocher, K. (2018). Can ISO-defined urban sustainability indicators be derived from remote sensing: An expert weighting approach. *Sustainability*, *10*(4), Article 1268. https://doi.org/10.3390/su10041268

López Chao, A., Casares Gallego, A., Lopez-Chao, V., & Alvarellos, A. (2020). Indicators framework for sustainable urban design. *Atmosphere*, *11*(11), Article 1143. https://doi.org/10.3390/atmos11111143

Lynch, A. J., Andreason, S., Eisenman, T., Robinson, J., Steif, K., & Birch, E. L. (2011). *Sustainable urban development indicators*. Penn Institute for Urban Research.

Merino-Saum, A., Halla, P., Superti, V., Boesch, A., & Binder, C. R. (2020). Indicators for urban sustainability: Key lessons from a systematic analysis of 67 measurement initiatives. *Ecological Indicators*, *119*, Article 106879. https://doi.org/10.1016/j.ecolind.2020.106879

Michalina, D., Mederly, P., Diefenbacher, H., & Held, B. (2021). Sustainable urban development: A review of urban sustainability indicator frameworks. *Sustainability*, *13*(16), Article 9348. https://doi.org/10.3390/su13169348

Miller, H. J., Witlox, F., & Tribby, C. P. (2013). Developing context-sensitive livability indicators for

transportation planning: A measurement framework. *Journal of Transport Geography*, *26*, 51–64. https://doi.org/10.1016/j.jtrangeo.2012.08.007

Mori, K., & Christodoulou, A. (2012). Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI). *Environmental Impact Assessment Review*, *32*(1), 94–106. https://doi.org/10.1016/j.eiar.2011.06.001

Nagy, J. A., Benedek, J., & Ivan, K. (2018). Measuring Sustainable Development Goals at a local level: A case of a metropolitan area in Romania. *Sustainability*, *10*(11), Article 3962. https://doi.org/10.3390/su10113962

Namavar, M., Moghaddam, M. R. A., & Shafiei, M. (2023). Developing an indicator-based assessment framework for assessing the sustainability of urban water management. *Sustainable Production and Consumption*, *40*, 1–12. https://doi.org/10.1016/j.spc.2023.06.006

OECD. (2021). *Applying evaluation criteria thoughtfully*. https://doi.org/10.1787/543e84ed-en

Ortiz-Moya, F., & Reggiani, M. (2023). Contributions of the voluntary local review process to policy integration: Evidence from frontrunner cities. *npj Urban Sustainability*, *3*, Article 22. https://doi.org/10.1038/s42949-023-00101-4

Papageorgiou, A., Björklund, A., & Sinha, R. (2025). A novel assessment framework for circular urban areas with its application to a Swedish municipality. *Sustainable Production and Consumption*, *57*, 121–137. https://doi.org/10.1016/j.spc.2025.05.005

Rathore, M. M., Ahmad, A., Paul, A., & Rho, S. (2016). Urban planning and building smart cities based on the Internet of Things using Big Data analytics. *Computer Networks*, *101*, 63–80. https://doi.org/10.1016/j.comnet.2015.12.023

Roche, C. (1999). *Impact assessment for development agencies: Learning to value change*. Oxfam.

Rodrigues, M., & Franco, M. (2020). Measuring the urban sustainable development in cities through a composite index: The case of Portugal. *Sustainable Development*, *28*(4), 507–520. https://doi.org/10.1002/sd.2005

Ruan, F.-L., & Yan, L. (2022). Challenges facing indicators to become a universal language for sustainable urban development. *Sustainable Development*, *30*(1), 41–57. https://doi.org/10.1002/sd.2227

Schetke, S., Haase, D., & Kötter, T. (2012). Towards sustainable settlement growth: A new multi-criteria assessment for implementing environmental targets into strategic urban planning. *Environmental Impact Assessment Review*, *32*(1), 195–210. https://doi.org/10.1016/j.eiar.2011.08.008

Schiavo-Campo, S., & Tommasi, D. (1999). *Managing government expenditure*. Asian Development Bank.

Selvik, J. T., Bansal, S., & Abrahamsen, E. B. (2021). On the use of criteria based on the SMART acronym to assess quality of performance indicators for safety management in process industries. *Journal of Loss Prevention in the Process Industries*, *70*, Article 104392. https://doi.org/10.1016/j.jlp.2021.104392

Shi, Y., Ge, X., Yuan, X., Wang, Q., Kellett, J., Li, F., & Ba, K. (2019). An integrated indicator system and evaluation model for regional sustainable development. *Sustainability*, *11*(7), Article 2183. https://doi.org/10.3390/su11072183

Simon, D., Arfvidsson, H., Anand, G., Bazaz, A., Fenna, G., Foster, K., Jain, G., Hansson, S., Evans, L. M., Moodley, N., Nyambuga, C., Oloko, M., Ombara, D. C., Patel, Z., Perry, B., Primo, N., Revi, A., Van Niekerk, B., Wharton, A., & Wright, C. (2015). Developing and testing the Urban Sustainable Development Goal's targets and indicators—A five-city study. *Environment & Urbanization*, *28*(1), 49–63. https://doi.org/10.1177/0956247815619865

Stitteneder, T. (2025). Wirkung wissenschaftlich analysieren: Methoden zur Evaluation sozialer Innovationen und gemeinwohlorientierter Organisationen. *ifo Schnelldienst*, *78*(2), 39–44. https://hdl.handle.net/10419/314785

Tanguay, G. A., Rajaonson, J., Lefebvre, J.-F., & Lanoie, P. (2010). Measuring the sustainability of cities: An analysis of the use of local indicators. *Ecological Indicators*, *10*(2), 407–418. https://doi.org/10.1016/j.ecolind.2009.07.013

Tran, L. (2016). An interactive method to select a set of sustainable urban development indicators. *Ecological Indicators*, *61*, 418–427. https://doi.org/10.1016/j.ecolind.2015.09.043

Turcu, C. (2013). Re-thinking sustainability indicators: Local perspectives of urban sustainability. *Journal of Environmental Planning and Management*, *56*(5), 695–719. https://doi.org/10.1080/09640568.2012.698984

UK National Audit Office. (2001). *Choosing the right FABRIC: A framework for performance information*.

UN-Habitat. (2012). *Input to post-2015 UN Development Agenda: Working Group B*. https://www.un.org/en/development/desa/policy/untaskteam_undf/groupb_unhabitat_suscities.pdf

Valencia, S. C., Simon, D., Croese, S., Nordqvist, J., Oloko, M., Sharma, T., Taylor Buck, N., & Versace, I. (2019). Adapting the Sustainable Development Goals and the New Urban Agenda to the city level: Initial reflections from a comparative research project. *International Journal of Urban Sustainable Development*, *11*(1), 4–23. https://doi.org/10.1080/19463138.2019.1573172

Verma, P., & Raghubanshi, A. S. (2018). Urban sustainability indicators: Challenges and opportunities. *Ecological Indicators*, *93*, 282–291. https://doi.org/10.1016/j.ecolind.2018.05.007

Villanueva, P. S. (2011). *Learning to ADAPT: Monitoring and evaluation approaches in climate change adaptation and disaster risk reduction—Challenges, gaps and ways forward* (SCR Discussion Paper 9). IDS. https://www.ids.ac.uk/download.php?file=files/dmfile/SilvaVillanueva_2012_Learning-to-ADAPTDP92.pdf

Webb, D., & Ayyub, B. M. (2017). Sustainability quantification and valuation. I: Definitions, metrics, and valuations for decision making. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *3*(3). https://doi.org/10.1061/AJRUA6.0000893

Zall Kusek, J., & Rist, R. C. (2004). *Ten steps to a results-based monitoring and evaluation system: A handbook for development practitioners*. The World Bank. http://hdl.handle.net/10986/14926

About the Authors

Oliver Peters (MSc) is a sustainability researcher at the German Institute of Urban Affairs, a PhD candidate at the Brandenburg University of Technology, and a lecturer of ecological economics at several universities. He leads projects on sustainable urban development with a focus on SDG indicators, reporting, and finance, as well as decentralised cooperation.

Valeska Liedloff (MA) is a researcher at the German Institute of Urban Affairs focusing on municipal circular economy, sustainability management, and indicator development for social-ecological transformation, combining quantitative and qualitative methods to support strategic sustainability frameworks and digital reporting tools aligned with global agendas.