



Understanding Frameworking for Smart and Sustainable City Development: A configurational approach

Organization Studies
1–22

© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/01708406221099694

www.egosnet.org/os



Leroy White

University of Exeter, UK

Katharina Burger 

University of Bristol, UK

Abstract

In recent decades, frameworks combining rankings and indices for smart and sustainable city development have proliferated. Stakeholders respond to them in various ways for strategizing towards urban sustainability. We refer to this as frameworking, which we identify as focusing on how frameworks are commensurated. However, research on commensuration has concentrated mostly on reactivity towards metrics. Little is known about how stakeholders contemplate the quality of and reaction to rankings and indices. We examine this issue through a configurational analysis of a set of European cities that consistently appear in these frameworks. We unveil several configurations of smart city metrics that relate to sustainability. Based on these effects, we theorize frameworking as differences in the relative configurations of smart city metrics that can generate performance. These configurations relate to three underlying dimensions: smart city capability, reactivity and context. We show that when frameworking is studied configurationally, we can identify the previously under-researched response to the quality of indices and reactivity to metrics. Finally, we discuss the theoretical and practical implications of a complex account of frameworks relevant to boosting urban sustainability.

Keywords

facticity, qualitative comparative analysis, rankings, reactivity, sustainability

Introduction

Interest in smart cities indices and rankings has grown significantly over the last decade (Mora, Bolici, & Deakin, 2017). These metrics are important to investors when allocating resources to smart city initiatives (Paroutis, Bennett, & Heracleous, 2014). Stakeholders see performance on

Corresponding author:

Katharina Burger, Faculty of Social Sciences and Law, University of Bristol, 11-13 Tyndall's Park Road, Bristol, BS8 1PY, UK.

Email: katharina.burger@bristol.ac.uk

them as vital for their strategic intent and reputation in competing for investments, advancing urban development and providing better services for citizens (Leydesdorff & Deakin, 2011). Meanwhile, scholars noted that they encourage cities ‘to pay attention to each other’ (Acuto, Pejic, & Briggs, 2021, p. 363), thus changing how cities compete (Kornberger & Carter, 2010). Furthermore, their relevance is amplified by the growing belief that they support finding solutions to intractable economic, social and environmental problems (Mora et al., 2017). Indeed, the interest in them is in part leading to their increasing multiplicity.

However, while smart cities indices and rankings are considered important, there are growing concerns about their proliferation (Acuto et al., 2021). In the quest for abridging the variety of metrics, scholars and stakeholders look to combine them into multi-dimensional frameworks so that the complexity inherent in smart cities appears simplified. Moreover, there are questions about the quality of metrics, where each metric used in frameworks is developed with a specific value system and its own standards of evidence (Giffinger, Haindlmaier, & Kramer, 2010). Therefore, while scholarship continues to enlighten about how these frameworks may lead to a better understanding of smart city development (Appio, Lima, & Paroutis, 2019), little is known about what distinctive effects arise from them, and their construction is rarely questioned (Mora et al., 2020). More knowledge is needed on how frameworks undergird smart city strategizing (Appio et al., 2019), how they engender competition between cities (Kornberger & Carter, 2010) and how they can seemingly demonstrate achievements towards progressive performance objectives (Acuto et al., 2021). This is alongside other studies that question the very idea of developing frameworks and metrics for smart and sustainable city development (Hollands, 2015). Scholars are also beginning to criticize studies on frameworks for inhibiting the development of a convincing theory from which an understanding of smart city performance can be derived (e.g. Mora et al., 2020). Thus, against this backdrop, research is warranted to forward an understanding of not just what it takes for cities to become smart and sustainable (Appio et al., 2019) but of how and why frameworks that use or combine metrics may make city strategizing for this endeavour possible (Kornberger & Carter, 2010).

However, even after several attempts at making a case for studying city strategizing (e.g. Czarniawska, 2002), organizational scholarship has yet to focus on this potential. Our goal is to address this concern. We do so by first drawing on organizational research on commensuration (Espeland & Sauder, 2007), which focuses on ‘the transformation of different qualities according to a common metric’ (Espeland & Stevens, 1998, p. 314), and whether this alters the way stakeholders strategize and think about their organization (Mazmanian & Beckman, 2018). Second, we attend to the issue that an organizational perspective on smart and sustainable cities is hindered by extant studies failing to recognize smart city settings as ‘problems of organized complexity’ (Patorniti, Stevens, & Salmon, 2018). Therefore, we suggest that these concerns can be remedied by attending to causal complexity (Ragin, 2000), that is, paying attention to the ways that metrics combine towards performance outcomes (Misangyi et al., 2016). To do this, we apply fuzzy set qualitative comparative analysis (fsQCA) (cf. Fiss, 2011) to a set of European cities that consistently appear in smart city indices and rankings. Thus, we propose that metrics combine in a configurational way (White, Lockett, Currie, & Hayton, 2021). We find that in applying fsQCA, several configurations lead to high or low performance towards the Sustainable Development Goals (SDGs) and quality of life (QoL). Moreover, we examine the different configurations identified from our analysis more closely and find that each configuration corresponds to varying combinations of metrics, inferring how progress towards smart and sustainable city development is being performed.

This study makes several contributions. First, we show how attention to a particular form of commensuration, frameworking, which is hitherto under-researched, opens a space to examine the

implications for settings with multiple metrics. We believe that a focus on frameworking will respond not only to calls to gain valuable knowledge about multiple metrics in the specific realm of smart cities and their performance for sustainable development (Mora et al., 2020) but also to calls to gain a deeper understanding of the impact of the commensuration of metrics more generally (Espeland & Sauder, 2007). Second, by adopting a configurational approach, we show how each combination takes on different meanings concerning competition and performance towards sustainability. Finally, we provide practical implications for stakeholders and scholars seeking to enumerate new frameworks and address the role of metrics more clearly as cities are strategizing to orientate towards sustainability.

Theoretical Considerations

We identify frameworking as a complex phenomenon depicting the dynamism around enumerating frameworks from different metrics, including measurements, indices and rankings and how they take on meaning in use, particularly concerning societal concerns such as sustainability (Appio et al., 2019). There are examples of this practice in several other settings, including education (e.g. Gunn, 2018), accounting (e.g. Pollock, D'Adderio, Williams, & Leforestier, 2018), and corporate social responsibility (e.g. Bermiss, Zajac, & King, 2014). In many of these studies, frameworking is predominantly a cognitive process. However, very few studies recognize that the power of frameworking is tied to how frameworks are produced and given authority (Mazmanian & Beckman, 2018). As such, we see that enumerating frameworks involves commensuration (Espeland & Sauder, 2007).

The advancement of commensuration into organizational life is often seen in terms of the ease with which the aggregation of incommensurable categories can be turned into metrics (Espeland & Stevens, 1998), that is, 'grouping them in the same frame, establishing original relations between them' (Callon & Muniesa, 2005, p. 1232). As such, commensuration requires significant resources, processes and methods, with actors and organizations integrating and reconfiguring different numbers and measures to establish novel interpretive frameworks (Espeland & Stevens, 1998). Most relevant for our study is scholarship that points to how metrics commensurate markets for societal concerns, raise concerns about their transparency and evoke reactivity to rankings as a way to infer that entities are competing. We review these in turn.

First, commensuration has been noted as offering the opportunity to create markets for social concerns (Huault & Rainelli-Weiss, 2011). For example, Sharkey and Bromley (2015), drawing on market information regimes theory, showed how the presence of metrics increases the salience and attention to particular societal issues, for example, pollution (Levin & Espeland, 2002). Further studies suggest that metrics boost domains as ones that can legitimately evaluate organizations according to concerns, such as environmental protection (Clementino & Perkins, 2020) or social responsibility (e.g. Slager, Gond, & Crilly, 2021). These studies show how commensuration conceives societal concerns as composed of metrics, for example, increased interest in metrics to advance governance, responsibility, and accountability as vectors of new public management (Meijer, 2018). However, the ambiguous notion of many social issues has led to increased interest in different metrics, all vying to be the means for explaining performance on these issues (Mennicken & Espeland, 2019). Furthermore, as stakeholders demand more accountability on societal concerns from organizations and governments, the proliferation of metrics and frameworks focusing on these interests seems unhindered (Espeland & Sauder, 2007). These effects suggest that it is still unclear whether commensuration induces markets that can focus precisely on particular societal concerns or whether competing on social issues is an unintended consequence of commensuration in that

they lead to metrics that are precarious and at risk of marketization (Mennicken & Espeland, 2019). This equivocality may be hindering attempts to further an understanding of the relationship between metrics and performance on social concerns, accountability and governance (Levin & Espeland, 2002).

Second, commensuration is as much a concern for the objective construction of metrics as it is that metrics are subjectively determined yet treated as an objective reality (Rao, 1994). With the former, a realist view prevails, which suggests that objective measures are necessary to understand what drives performance (i.e. their predictive power). Indeed, stakeholders are drawn to objective indices of complex reality. Here, producers vehemently defend that metrics commensurate perceived objectivity and emphasize that indices need to be transparent to reduce uncertainty, ambiguity and misunderstandings (Espeland & Stevens, 1998). Transparency (Mennicken & Espeland, 2019) is a drive to establish a mindset for precision and exactness (Rao, 1994) and for metrics to meet a ‘standard of desirability’ (Graffin & Ward, 2010, p. 331). On the other hand, other scholars point out that the quest for transparency is impossible (Espeland & Sauder, 2007), which brings doubts about the soundness of metrics (e.g. Roberts, 2018). Taking these different views on transparency together, it remains unclear whether metrics can ever be a good signal or gauge of performance, or both (Brunsson, Rasche, & Seidl, 2012). This is not to deny that metrics need to be transparent. Scholars show that despite uncertainty about their transparency, stakeholders often come to make sense of metrics as taken-for-granted facts and use them to inform future strategic actions (Bermiss et al., 2014). Indeed, when the quality of a metric is assumed or challenged, its facticity (the quality of taken-for-grantedness) comes to the fore (Power, 2021). This is to ensure that metrics do not only represent purely pre-given objective details about organizational performance but also construct what comes to be recognized as performance (Mazmanian & Beckman, 2018). However, seeing both perspectives together to study their implications remains an ongoing challenge (Power, 2021).

Finally, most studies on commensuration focus on the reactivity that metrics invoke in that they pressure organizations to conform and perform regarding how they compare relative to others (Espeland & Sauder, 2007). Reactivity is defined as changing ‘behaviour in reaction to being evaluated, observed, or measured’ (Espeland & Sauder, 2007, p. 1). For example, many studies show how universities are assessed according to rankings and adopt strategies aligned to the values embodied in them (Marginson & Van der Wende, 2007).

Reactivity is seen by some as good for the field; for instance, motivating poorly ranked organizations to improve their performance (Chatterji & Toffel, 2010) or encouraging organizations to raise their ambition by continually comparing their performance to their peers (Rowley, Shipilov, & Greve, 2017). This is regarded as striving to enhance positional status (Sauder & Espeland, 2009). However, rankings are also criticized for causing anxiety and increasing organizational resistance to them (Gerdin & Englund, 2019). It is also suggested that organizations may develop ambivalence towards them (Sauder & Espeland, 2009). Consequently, practical and sceptical responses can occur, potentially leading to means–ends decoupling (Slager & Gond, 2022) as organizations engage in activities as a result of their performance on a metric that are weakly linked to their goals and turn out to be largely ineffective (Wijen, 2014). Thus, there are many ways stakeholders react to rankings, leading some scholars to suggest that it is vital to consider the multiple facets of reactivity to rankings and examine the different ways they affect organizational behaviour (Pollock et al., 2018).

Some scholars increasingly suggest that it is important to revisit assumptions about reactivity to rankings in settings with multiple metrics (Pollock et al., 2018). It is also not entirely clear how reactivity to rankings in settings with multiple metrics affects organizational behaviour (Bermiss et al., 2014). Indeed, it is claimed that settings with various metrics will weaken reactivity to

rankings due to the broad uncertainty and variation regarding their quality (Kim, 2020). Others highlight that numerous metrics in a setting may raise questions about which one to align to (Pollock et al., 2018). Further, some show that reacting evenly and broadly to multiple metrics would mean being pulled in different directions.

In summary, once a metric or framework is enumerated, it can raise the salience of societal concerns and be valued as a signal of performance and may be a facet of an organization's competitive outlook. The focus of many past studies on commensuration is on a single issue (either transparency or reactivity), which appears insufficient. As such, it seems essential to embrace different issues together in that this may highlight how they capture complex organizational processes (Katz & Kahn, 1966). By studying the issues together, we can close the 'complexity differential' (Schneider, Wickert, & Martí, 2017, p. 183), to better understand the issues grounded in complex settings. This would require exploring the complexities inherent in the relationships among organizational phenomena.

Thus, to advance an understanding of the complexities of frameworking, we propose a configurational approach. This is important because we assume that metrics representing organizational phenomena 'combine into distinct configurations to produce an outcome of interest' (Misangyi et al., 2016, p. 257). Indeed, a configurational perspective has long been associated with organizational design and effectiveness (Doty, Glick, & Huber, 1993). We therefore adopt this view to consider the complex interactions of metrics and their implications for performance. Moreover, it has been highlighted that scholarship on smart cities' performance is hindered because 'causal agency and mechanisms are not theorized' (Mora et al., 2020, p. 5). Thus, a configurational approach is promising because it can infer these characteristics via the notion of causal complexity (Ragin, 2000). Here, causation is seen as multifaceted in character (Misangyi et al., 2016), which requires attention to the many ways a 'common outcome is reached' (Ragin 2000, p. 88), that is, equifinality. Causal complexity also allows for the possibility that relationships 'causally related in one configuration may be unrelated or even inversely related in another' (Meyer, Tsui, & Hinings, 1993, p. 1178), that is, asymmetry. A configurational approach also views causal relations and mechanisms as conjunctural (Fiss, 2011). Scholars maintain that a conjunctural perspective can be used to examine causal complexity in that it seeks ways to theorize the connections between concepts to forward a more integrated understanding of organizational phenomena (Schneider & Wagemann, 2012).

Empirically, causal complexity and conjunctural mechanisms can be investigated through the logic of set theory (Ragin, 2000). This allows us to conceptualize configurations and their conjunctural nature as combinations of fuzzy sets (Fiss, 2011). These sets can be used to account for the fuzziness in the commensuration of rankings or indices and study the consequences of frameworks with multiple metrics (Ragin, 2000). For instance, rather than focus on rankings as a metric that establishes only the position of organizations relative to each other, fuzzy sets replace this view with a more nuanced one that looks at the degree of membership in well-defined sets. Attention to fuzzy sets involves familiarity and flexibility as the researcher learns more of the metrics, their production and their instances (Ragin, 2000). As such, a fuzzy set approach seems particularly useful for conceptualizing frameworking and commensuration in our study (Ragin, 2000).

Methods

Sample and data

In this study, we bring together data from several smart city indices and rankings.¹ The indices and rankings were selected based on our theoretical interest and prior research on smart cities (cf.

Appio et al., 2019). These metrics are produced from public and private sources. We include indices and rankings that provide detailed methodological information and full access to the raw data sources (cf. Sharifi, 2020). Details of the indices and rankings are provided in Table 1.

Identifying a meaningful set of cities to include in our study is non-trivial, noting there may be bias in city selection in smart city rankings and that a city can be ranked differently in different indices (Meijering, Kern, & Tobi, 2014). Our selection of the cities is based on the following. First, we limit the case pool to cities within the European Union (EU) because the EU constitutes an influential policy background that helps shape smart city development, for example, through significant funding under EU cohesion policy (Mora & Deakin, 2019). Second, we rely in our choice on city size as a factor to select a relatively homogeneous group to prevent our analysis from ‘comparing apples and oranges’ (cf. Meijering et al., 2014). We also choose large cities² because they are important in our consideration of the start-up, entrepreneurial scene, which is unlikely to be present to the same extent in medium and smaller cities (Ivaldi, Penco, Isola, & Musso, 2020). Third, we consider the question of whether the cities would constitute a meaningful group of competitors. Scholars note that metrics mediate competition not solely as the product of managerial cognition (Porac & Thomas, 1990) nor just through audience interpretation (Pollock et al., 2018). Thus, our assumption is built on the idea that the effect of being included in rankings is that cities may perceive they are competing. For instance, Kornberger and Carter (2010) highlighted how rankings lead to perceived competition between cities, for example London and Sydney, where none had existed previously (Kornberger & Carter, 2010). Therefore, the presence of our sample cities in multiple rankings and indices was considered as potentially giving rise to rivalry effects. Finally, we arrived at the following cities as cases in our final sample: Barcelona, Berlin, Bratislava, Brussels, Budapest, Copenhagen, Dublin, Hamburg, Helsinki, Lisbon, London, Madrid, Munich, Paris, Prague, Rome, Stockholm, Tallinn and Vienna. For the fsQCA, we believe the sample is robust because the sample cities balance coverage with complexity.

Conditions and outcomes

FsQCA requires that each city is a case that can be examined as a set of conditions. We relied upon case knowledge and theoretical insights to derive our set of conditions. While multiple metrics and frameworks coexist that seek to capture conditions for smart and sustainable city transitions, three of our conditions align with what others have referred to as ‘foundations’ (cf. Yigitcanlar, Han, Kamruzzaman, Ioppolo, & Sabatini-Marques, 2019), in that smart cities require the alignment of technical (SMART), social (OPEN) and institutional (QoG) elements. The inclusion of regional innovation capacity (RIC) was informed by prior work on smart city performance, which included references to regional innovation systems. The decision to include entrepreneurial support (MENT) was based on prior research that claims that this dimension is often attributed to smart city development. Lastly, we included the Cities in Motion Index (REP) as a condition. It is a well-established ranking and is noted by stakeholders and the media, as well as other users of the ranking. Thus, any change in the rankings becomes noticeable as evoking reactivity.

We also identified two main outcome measures, recognized as desirable and integrative measures of performance towards sustainability: the SDGs and QoL (Ivaldi et al., 2020). For the robustness check, we included a measure of gross value added (GVA). An overview of the conditions and outcomes is provided in Table 1.

Calibration

An important step before any fsQCA analysis is calibrating each of the conditions and outcomes into fuzzy sets. Calibration involves setting thresholds for each condition (Ragin 2000), and we

Table 1. Data sources.

Condition	Data source	Reference	Data processing	Year	Use
Regional Innovation Capacity (RIC)	Eurostat	European Regional Competitiveness Index	z-scores	2019	Business efficiency, science and research knowledge spillovers, and general technological infrastructure
Quality of Government (QoG)	Citizen survey data	Smart City Index, Institute for Management Development	Average % of respondents who agree or strongly agree with each item	2019/2020	Citizen engagement, participation in decision-making, and accessibility of information from the local government
Openness (OPEN)	Eurostat	European Cultural and Creative Cities Monitor	Scaled scores (0 to 100), 100 is highest	2019	Openness, tolerance and trust
Smart technology (SMART)	Unique data sources on city-level digitalization and smart services	Easypark Smart City Index	1 to 10; 10 is highest	2019	Transport and mobility, governance, innovation economy, digitalization, and expert perception
Entrepreneurial Support (MENT)	Data from multiple sources, including Eurostat	European Digital City Index	0 to 1; 1 is highest	2016	Networking and mentoring events, access to accelerators, and public sector information with relevance to digital technologies
Reputation (REP)	Data from multiple sources	IESE CIMI (Cities in Motion) Index, IESE Business School	Rank scores for 2016 and 2019	2016–2019	Reputation results from collective sensemaking by producers and users of rankings (Fombrun & Shanley, 1990). We also build on the importance of the notion of position in a ranking, but we focus on difference in that, following Kim (2020), we note that changes in differences in rank are salient.
Sustainable Development Goals (SDG)	Data from multiple sources, including Eurostat	Sustainable Development Solutions Network	0–100; 100 is highest	2019	UN 17 Sustainable Development Goals (SDGs)
Quality of Life (QoL)	Unique crowdsourced data	Numbeo is a crowdsourced global database	Geo-filter 'Europe'. A higher score is better	2019	Purchasing power index, pollution index, house price to income ratio, cost of living index, safety index, health care index, traffic commute time index and climate index (lowest: Rome, 108.58; highest: The Hague 194.52)
GVA per worker (GVA)	OECD regional data	OECD	z-scores, excluding non-regionalized data	2018	Contribution to the economy of each individual

GVA, Gross value added; OECD, Organisation for Economic Co-operation and Development; IESE, Instituto de Estudios Superiores de la Empresa/ Institute of Higher Business Studies; UN, United Nations.

heed the point that this process must involve substantive and theoretical knowledge. Fuzzy sets for our conditions and outcomes are calibrated and assigned values between 0 and 1. We follow the general practice of calibrating our measures on a three-way basis. Here, we define full membership as 1, non-membership as 0 and the crossover point as 0.5 (the point of maximum ambiguity, i.e. where we cannot tell whether the case is more in or out of the set). Thus, our datasets are calibrated by the direct method (Ragin, 2000). Calibration allows the uncovering of greater granularity in the measures. Therefore, we could combine substantive case knowledge (e.g. the enumeration methods used for the metrics) and an understanding of the data. Furthermore, in the absence of an external benchmark, we used technical criteria such as the distribution of cases. For instance, we considered that even though we only included data for 19 cities as cases for the analysis, the raw data are from larger samples, for example, OPEN (190 cities); therefore, we used the distributions to decide set membership. Full details of the calibration of our conditions are provided in Table 2.

Calibration is also how we account for commensuration, including our interest in transparency and reactivity. By assessing the degree of membership, we focus on the transparency of the metric, and by calibrating according to theory and substantive knowledge, we focus on the reactivity to rankings. For instance, both the SMART and MENT conditions are based on indices (see Table 1). The perceived objectivity or taken-for-grantedness by stakeholders is difficult to assess directly from the measure. Based on detailed knowledge of indices, we appraise the measures regarding their degree of membership in the conceptual category. We defined a strong influence ('fully in' = 1) when the condition is fully externalized and objectified. We defined cases as a weak influence ('fully out' = 0), where stakeholders' reflection on transparency is high. These corresponded to the 75th and 25th percentiles in the indices, respectively. We were also confident in setting the crossover point at the median (see Table 2). Our condition REP is concerned with reactivity to the Cities in Motion Index, which ranges from a substantial rise, no change, or a fall in the ranking, representing the idea that cities react to the rankings and enact actions in a way that aligns with a metric's criteria (see Table 2). Finally, as a team, we worked through the data and agreed on the final calibrations by adopting a triangulation approach (Schneider & Wagemann, 2012) to reduce researcher bias.

Empirical analysis

The analyses for this study were performed with fsQCA 2.0 (Ragin, Drass, & Davey, 2006). First, we checked whether any of our conditions were exclusive enough to cause any of the outcomes. This is a test for necessary conditions that indicate whether the conditions play a role in isolation in producing the outcomes. In our data, none of the conditions was found to be necessary for any of the outcomes (SDG, QoL, GVA). The finding indicates that the interrelations between the conditions are complex processes that are configurational. We approach this issue empirically by examining the configurations in terms of causal sufficiency.

Our sufficiency analysis involves constructing a truth table with 2^k rows, where k is the number of causal conditions used in the analysis, then logically reducing the table based on a frequency cut-off threshold (the number of cases to be considered) and the consistency of the solution. It is up to the researcher to set a priori the threshold for the frequency of cases per configuration and the minimum thresholds for consistency. We set the suitable frequency for the cut-off to one case and the consistency at 0.8 (Ragin, 2000). This gave us over 85% of cases to include in the analysis for all our outcomes. Following this, we used the natural break in raw consistency scores as the threshold consistency and reduced the solutions using the software's Quine–McCluskey algorithm. The software calculates the raw coverage and the unique coverage as measures showing empirical importance, indicating the extent to which the outcome is explained by a causal condition. Raw

Table 2. Calibration of conditions.

Condition	Calibration	Fuzzy Score	Exemplar cities
RIC	High innovation capacity (greater than 1) is fully in Crossover point on this measure is 0.65 Low innovation capacity (lower than -0.04) is fully out	1 0.5 0	London, Stockholm, Copenhagen, Helsinki Budapest, Barcelona, Rome, Tallinn
QoG	High-quality local governance (greater than 57.65) is fully in Crossover point on this measure is 54.7% Low-quality local governance (lower than 47.3%) is fully out	1 0.5 0	Copenhagen, Munich Budapest, Lisbon
OPEN	High degree of openness (greater than 41) is fully in Crossover point on this measure is 33 Low degree of openness (lower than 17) is fully out	1 0.5 0	London, Munich Prague, Bratislava, Rome
SMART	High presence of smart technology (greater than 6.7) is fully in Crossover point on this measure is 6 Low presence of smart technology (less than 5.8) is fully out	1 0.5 0	Tallinn, Helsinki Prague, Rome
MENT	High level of entrepreneurial support (greater than 0.5) is fully in Crossover point on this measure is 0.3 Low level of entrepreneurial support (less than 0.1) is fully out	1 0.5 0	Dublin, Helsinki, London Paris, Vienna, Copenhagen Prague, Brussels, Hamburg, Rome
REP	High striving cities (greater than 10) is fully in Crossover point on this measure is 0 Non-striving cities (less than -5) is fully out	1 0.5 0	Stockholm, Rome London, Paris Budapest, Brussels
SDG	Cities making good progress to achieving all SDG goals (64.3%) is fully in Crossover point on this measure is 60% Cities making poor progress (less than 56%) is fully out	1 0.5 0	Tallinn, Helsinki Prague, Rome
QoL	High quality of life (greater than 170) is fully in Crossover point on this measure is 155 Low quality of life (less than 120) is fully out	1 0.5 0	Vienna, Copenhagen, Helsinki Paris, Milan, Rome
GVA	High relative GVA (greater than 1.5) is fully in Crossover point on this measure is 1 Low relative GVA (less than 0.1) is fully out	1 0.5 0	Dublin, Paris, London, Stockholm Lisbon, Tallin, Budapest

coverage shows the proportion of cities that are ‘fully in’ the present conditions of a particular configuration and unique coverage shows the proportion of cities covered uniquely by a particular solution (Schneider & Wagemann, 2012, p. 139). It should be noted that these coverage scores reflect the empirical strength that can be attributed to an individual configuration. We also reported the overall consistency scores that measure how likely each of the configurations leads to the outcome specified. Finally, we followed the convention established by Fiss (2011) to present our findings. In particular, ‘•’ denotes the presence of a condition, ‘o’ represents its absence, and a blank space indicates that a given condition is not causally related to the outcome. We followed the convention to denote larger circles to indicate core conditions (these are part of the parsimonious and intermediate solutions), whereas small circles refer to peripheral conditions that only occur in intermediate solutions. We present all the configurations found and the coverage and consistency scores in Tables 3 and 4.

To ensure the robustness of the results, the choices made for the calibration of the measures, the frequency thresholds and consistency levels are examined by running the analysis varying these choices (Schneider & Wagemann, 2012, p. 139). We checked several crossover points for our measures and found that changes had a marginal influence on the solutions. With the frequency threshold, we set the cut-off at 2, which led to a small reduction in the number of solutions generated. We also varied the consistency level to a more demanding level (0.85), which led to no change in the number of configurations. We conclude that we found no significant deviation from our presented results.

Findings

Model A1 (Table 3) exhibits three configurations that achieve high levels of the SDG outcome (solution consistency, 0.97; solution coverage, 0.64). This high solution consistency score strongly supports these three configurations in that the configurations are consistently associated with high performance towards the SDG outcome. All configurations in modelA1 are characterized by a combination of a high presence of openness, tolerance and trust (OPEN) and a high presence of SMART, leading to high levels of progress towards the SDGs. These are core conditions for the configurations. In addition, all three configurations are characterized by a high quality of local government (QoG).

Table 3. QCA results: High performance.

Condition	Model A1			Model A2			Model A3
	SDG1	SDG2	SDG3	QoL1	QoL2	QoL3	GVAI
RIC	•		•	•	○	●	●
SMART	●	●	●	•	•	●	•
OPEN	●	●	●	•	•	•	•
REP		•	•	○	•	•	●
QoG	•	•	•	●	●	●	•
MENT	•	•		•	•	○	
Raw coverage	0.53	0.45	0.45	0.54	0.34	0.34	0.38
Coverage	0.13	0.05	0.06	0.20	0.34	0.04	0.38
Consistency	0.97	0.98	0.98	0.97	0.95	0.98	0.92
Solution coverage	0.64			0.63			0.38
Solution consistency	0.97			0.96			0.92

The first configuration, SDG1, shows that a combination of all conditions, except for the absence of REP, is sufficient to generate high progress towards the SDGs (raw coverage, 0.53; consistency, 0.97). SDG2 uncovers a similar core scenario with a high level of REP and a high level of MENT (raw coverage, 0.45; consistency, 0.98). SDG3 shows how core conditions and QoG combined with a high level of RIC and with reputation (REP) are sufficient to generate high progress towards the SDGs (raw coverage, 0.45; consistency, 0.98).

Table 3 also exhibits three configurations leading to QoL (consistency 0.96; coverage 0.63). The high consistency score strongly supports these configurations. The configurations (modelA2) are characterized by a high QoG as a core condition leading to a high QoL. The configuration QoL1 shows that a combination of a high level of QoG and the absence of high levels of REP, as core conditions, together with a presence of all other conditions, lead to a high level of QoL (raw coverage, 0.54; consistency, 0.97). For QoL2, the presence of high levels of QoG and SMART and the absence of RIC and high levels of all other conditions characterize this configuration (raw coverage, 0.34; consistency, 0.95). QoL3 is characterized by the high presence of QoG and the presence of RIC, and the absence of MENT as core conditions (raw coverage, 0.34; consistency, 0.98).

We report another outcome condition, GVA, reported as modelA3 (Table 3). We use this model as a robustness check. A3 exhibits one configuration leading to the outcome (raw coverage, 0.38; consistency, 0.92). The configuration shows the core conditions of RIC and REP and high levels of all the other conditions, except MENT.

Using fsQCA, we explored which combinations of conditions lead to the absence of the outcomes (see Table 4). Here, we found evidence of asymmetric causality for low progress towards the outcomes (~SDG, ~QoL and ~GVA). The models' solution consistency scores were over the acceptable consistency threshold of 0.85. This is strong evidence for the configurations uncovered.

Model B1 in Table 4 reveals configurations for each of the outcomes. The first configuration is characterized by a high level of openness as a core condition, together with a strong reputation and entrepreneurial support, and the absence of all other conditions (RIC, SMART, QoG), leading to low levels of SDG (~SDG1; raw coverage, 0.21; consistency, 0.99). Two configurations leading to low levels of QoL are found: ~QoL1 (modelB2) is characterized by the presence of MENT, the absence of

Table 4. QCA results: Low performance.

Condition	Model B1	Model B2		Model B3	
	~SDG1	~QoL1	~QoL2	~GVA 1	~GVA2
RIC	○	○	○	○	○
SMART	○	○	○	○	○
OPEN	●	○	○	○	○
REP	●	○	●	○	●
QoG	○	○	○	○	○
MENT	●	●	○	●	
Raw coverage	0.21	0.15	0.16	0.23	0.23
Coverage	0.21	0.09	0.09	0.08	0.07
Consistency	0.99	0.90	0.90	0.95	0.97
Solution coverage	0.21	0.24		0.31	
Solution consistency	0.99	0.94		0.96	

SMART and REP (raw coverage, 0.15; solution consistency, 0.9), and the absence of all other conditions; \sim QoL2 (modelB3) (raw coverage, 0.16; consistency, 0.9) is characterized by high REP and the absence of QoG and MENT as core conditions, as well as the absence of all other conditions.

For the robustness check, the GVA outcome resulted in two configurations \sim GVA1 (raw coverage, 0.23; consistency, 0.95) and \sim GVA2 (raw coverage, 0.23; consistency, 0.97). These are identical in the absence of RIC, SMART, OPEN and QoG as core conditions.

Comparative multi-level analysis

We investigated how the contextual conditions influence performance. Therefore, we analysed the cases following the procedures used for comparative multi-level analysis (Denk & Lehtinen, 2014). The analysis is presented in Table 5.

When we examined the role of context, we found that the consistency scores of the configurations range from 0.75 to 0.99. This indicates that we can reasonably assume that the configurations are consistently associated with the outcomes, but there are contextual effects. For instance, we noticed that the contextual effects of openness (consistency, 0.98) and QoG (consistency, 0.98) on the conditions associated with performance towards SDG are higher than the effects of the full model (all cases; consistency, 0.96). However, there is a weaker effect of the absence of these conditions on SDG (i.e. the consistency measure is lower than the full model with all cases). Specifically, for openness, we noted here that the consistency score (consistency, 0.71) is lower than the critical value of 0.8 (cf. Ragin, 2000), indicating that the absence of openness is insufficient for the SDG outcome. Regarding the measure for innovation capacity, RIC, the consistency measure is lower than the full model with the outcome for the QoL (consistency, 0.84). In terms of the QoG, the consistency measure is lower than the full model with all cases of low performance on the QoL measure (consistency, 0.75). These results confirm differences in the causal effect of the contextual conditions on the outcome measures.

Table 5. QCA results: Contextual analysis.

			CONTEXT			
			RIC	OPEN	QoG	
OUTCOME	SDG	Consistency (coverage)	0.96 (0.53)	0.98 (0.41)	0.98 (0.4)	Present
			Helsinki	Stockholm	Stockholm	
			0.97 (0.17)	0.71 (0.01)	0.62(0.31)	
	\sim SDG	Consistency (coverage)	Vienna	Prague	Lisbon	Present
			0.99 (0.12)	0.99 (0.21)	0.99 (0.15)	
			Brussels	Budapest	Prague	
	QoL	Consistency (coverage)	0.99 (0.21)	0.99(0.18)	0.99 (0.2)	Absent
			Barcelona	Rome	Madrid	
			0.84 (0.17)	0.96 (0.30)	0.98 (0.18)	
	\sim QoL	Consistency (coverage)	Helsinki	Hamburg	Vienna	Present
			0.95 (0.15)	0.89 (0.21)	0.91(0.15)	
			Prague	Vienna	Lisbon	
\sim QoL	Consistency (coverage)	0.85 (0.1)	0.88 (0.18)	0.75 (0.11)	Present	
		Brussels	Barcelona	Budapest		
		0.93 (0.24)	0.93 (0.24)	0.93 (0.24)		
			Rome, Budapest	Rome	Rome	Absent

Discussion

Our findings show that important conditions leading to high performance towards the SDG outcome include the presence of high QoG, openness and smart technology, which are present in all three configurations (SDG1, SDG2 and SDG3). The joint presence of these conditions aligns with prior research that suggests that these conditions interact in urban settings (Leydesdorff & Deakin, 2011) and with research that suggests that sustainability benefits cannot be realized without openness together with high local government quality and smart city technologies, confirming that these are relevant for cities attempting to transition towards sustainability. We also observed differences in the configurations leading to positive performance towards achieving sustainability goals regarding the presence of subjective conditions (i.e. reactivity to a ranking of the cities).

Configuration SDG1 represents innovation approaches aligning with investments in human and social capital and high-quality ICT infrastructures, enabling progress towards the SDGs and QoL (Caragliu & Del Bo, 2019). We label this the ‘integrative’ configuration. This configuration accords with prior research that suggests that performance derives from an ability to balance or resolve contradictions from the conditions and context better than others (Gianelle, Guzzo, & Mieszkowski, 2020). Specifically, we see in this configuration the innovation capacity leveraged in tandem with entrepreneurial support, which confirms a distinct connection between sustainable city development and the fields of entrepreneurship and innovation (Leydesdorff & Deakin, 2011). Exemplars for this configuration include Helsinki, Copenhagen and London. These cities are repeatedly high performers in rankings. But we note the absence of reactivity to rankings for performance. The results do not fully support that high performers on the rankings are more likely to flaunt their position to external audiences (e.g. Elsbach & Kramer, 1996). Indeed, it is suggested that some high performers on rankings may remain ambivalent about such an achievement because they do not want to be accused of hypocrisy (e.g. Carlos & Lewis, 2018).

Configuration SDG2 shows that the presence of entrepreneurial support is a key condition relevant to performance. Therefore, we labelled this as the ‘entrepreneurial discovery’ configuration. This aligns with prior research, which showed that facilitating entrepreneurial activity to stimulate knowledge spillovers and innovation is regarded as central to smart development, driving innovations (Richter, Kraus, & Syrjä, 2015). Vienna is our exemplar for this configuration. It has been identified as a hotspot for entrepreneurs and is an example of a city that uses its sphere of influence to create an environment for entrepreneurial activity to unleash the potential of digital technologies to satisfy urban sustainability needs (Brandtner, Höllerer, Meyer, & Kornberger, 2016). This configuration also highlights the importance of reactivity to rankings as a reputational signal in association with the outcome. Again, Vienna is exemplary of here in that it is noted that its goal seems to be winning the international inter-city competition (Brandtner et al., 2016). This supports studies that confirm that it is not only socio-technical conditions that matter; the cities’ attractiveness is also important for the involvement of entrepreneurial individuals and organizations (Jessop & Sum, 2000). The findings also confirm studies that show the importance of positive changes in achievements in rankings, which can alter perceptions of reputation (Kim, 2020), in that rankings return reputational advantages, mostly to striving organizations (Schultz, Mouritsen, & Gabrielsen, 2001). Finally, configuration SDG3 suggests that in some cities, a high concentration of learning and innovation (Richter et al., 2015) is associated with sustainability. We labelled this the ‘innovation accomplishment’ configuration. The exemplars are Hamburg and Stockholm, which have been reported to have a diversified knowledge base. We also see the presence of striving in this configuration.

In addition, our analysis captured one configuration that leads to low performance towards the SDGs (~SDG1), characterized by the core condition of openness and the high presence of

entrepreneurial support, as well as a presence of striving in the rankings. However, this configuration also includes a low presence of innovation capacity (RIC), QoG and smart technology. In contrast to SDG1, the presence of openness and entrepreneurial support is still characteristic of cities working towards achieving the SDGs, but not sufficient within this configuration. Thus, we uncovered an interesting challenge to the central role attributed to openness in much prior work on city development. Moreover, with this configuration, striving in the rankings is not sufficient for progress towards the outcome. On the one hand, this aligns with extant studies showing that low performers are more reactive to rankings (e.g. Chatterji & Toffel, 2010) in that stakeholders will organize their practices to accommodate these metric requirements, which might lead to superficial compliance to the metrics. On the other hand, it confirms prior research that suggests that attention to rankings could lead to the neglect of attention to complex interactions in smart city development (Giffinger et al., 2010). This configuration elides with studies on rankings that argue that attention to rankings can lead to means–end decoupling (Espeland & Sauder, 2007), which is not just due to a lack of resources. The exemplar cities are Madrid and Barcelona, suggesting that this configuration describes a structurally and institutionally less well-resourced smart city setting. However, when linked to challenging settings, decoupling dynamics are more salient (Slager & Gond, 2022). Therefore, we labelled this configuration ‘weakly coupled’.

QoL constitutes another outcome that cities often pursue (De Guimarães, Severo, Felix Júnior, Da Costa, & Salmoria, 2020). The three identified configurations leading to high performance on the QoL outcome are characterized by a high QoG as a core condition. This confirms prior studies that show that the QoG interacts with other conditions to be sufficient for improving QoL (Simmons, Giraldo, Truong, & Palmer, 2018). These combinations of conditions match the labels we discovered for the SDG configurations, in that QoL1 is integrative, QoL2 an entrepreneurial discovery configuration and QoL3 an innovation accomplishment. QoL1, exemplified by Munich, seems to reflect an all-around combination of conditions leading to high performance with respect to the QoL. In comparison, QoL2 and QoL3 show the presence of a change in the rankings’ conditioning performance.

In terms of poor performance on the QoL measure, two configurations were found. Overall, these configurations reflect the low presence of all other conditions. We labelled them both as ‘fragmented’ configurations. \sim QoL1 shows that despite the presence of entrepreneurial support, the configuration leads to poor performance. This configuration is exemplified by Budapest, confirming prior research that the presence of a single condition is insufficient for performance in complex urban systems (Mora et al., 2020). \sim QoL2 illustrates that striving in the rankings together with the low presence of all other conditions leads to a low QoL, exemplified by Rome, which is similar to \sim SDG1.

In summary, our study identified and analysed five configurations (i.e. integrative, entrepreneurial discovery, innovation accomplishment, weakly coupled, and fragmented). We found that they display differences in their relation to progress towards the SDGs and QoL. We now turn to the importance of the relationship between contextual and smart city conditions for performance.

Table 5 highlights that openness (OPEN) and QoG are contextual conditions for SDG, with Stockholm being an exemplar. This finding aligns with prior research that, for example, highlights that in the pursuit of sustainability, openness is seen to lead to the development of multi-sectoral partnerships, governance and citizen trust that supports a high level of innovation (Yigitcanlar et al., 2019). High openness may also signal that a city is receptive to new ideas, which may stimulate experimentation and innovation (Sengers, Wieczorek, & Raven, 2019). Moreover, extant studies have shown that openness is critical for the attractiveness of a city to creative workers. Vienna

is an exemplar that actively sought to engage entrepreneurs by introducing subsidies for climate-friendly technologies (Brandtner et al., 2016).

Finally, in contrast, the absence of these contextual conditions (OPEN, RIC and QoG) is associated with poor performance on the outcomes. This accords with prior research that highlights that innovation can only occur in a supportive governance environment. At the same time, limited governance capacity is associated with a reduced ability of a city administration to stimulate private-sector demand for research and development and address the weak embedding of research and technological infrastructure, for which the absence of RIC is indicative. As such, the absence of the three contextual conditions suggests a decoupled environment where innovative smart city development practices are not facilitated. This is exemplified by Rome. Thus, overall, our findings confirm extant smart cities studies that claim that theories considering transitions to sustainability as multi-level-phenomena are better at understanding how context is expected to affect the relationship between the causes and the outcome (Geels, 2002), which is also important in understanding the mechanisms of urban sustainability (Mora et al., 2020).

Implications and Conclusion

The configurational nature of our findings supports the importance of examining the interactions of metrics and context together and in more detail. We demonstrate empirically that these relationships are more complex than typically assumed (Mora et al., 2020). In short, we show that the effects of metrics cannot be understood in isolation as assumed in prior research. However, many studies also combine metrics into frameworks, which are almost universally reliant on trade-off approaches (including methods such as cost-benefit analysis and multi-attribute decision making or additive methods such as linear regression analysis that assume causal relationships as the covariation between dimensions and outcome variables). On the other hand, our approach, drawing on set theory, allowed us to transcend data types and linear effects analysis (Abbott, 1988); therefore, we account for the multiplicity and interwovenness between metrics and outcomes lacking in extant studies (Mora et al., 2020). Furthermore, the approach aligns with our argument for causal complexity (Meyer et al., 1993), assessing causal relations as conjunctural and providing evidence for the sufficiency of distinct configurations for the outcome of interest (Misangyi et al., 2016).

Table 6 summarizes the result from the fsQCA. Our findings show clear differences in the relative configurations of smart city metrics that can generate performance, indicating equifinality. These configurations also relate to three underlying dimensions: smart city capability (strong or weak influence), reactivity (ambivalence or striving to improve) and context (narrow or broad). Each dimension has different implications for the progress towards sustainability and QoL. However, we also found a thought-provoking result in that the presence of a condition associated with the dimension of context (e.g. openness) is found for both high and low performance (e.g. towards the SDGs). Therefore, we also found differences in configurations leading to high and low performance of the outcome measures, showing asymmetry.

Moving on to discuss the implications of our work, our study addresses scholars' continued concerns around the theoretical and practical equivocality in explaining smart city developments towards progressive goals such as urban sustainability (Mora et al., 2020). By integrating insights from scholarship on commensuration, and configurational and conjunctural perspectives, we clarify some of the issues raised in the literature. We demonstrate this clarity by expanding on the dimensions advanced in our study (Table 6).

Our study considered metrics as signals of smart city capability, recognizing that the extent to which metrics can be used to understand the performance of cities remains uncertain

Table 6. A framework of performance towards sustainability.

Configurations	Impression of smart city capability		Reactivity to rankings	Context	Outcome	
	SMART	MENT	REP		QoL	SDG
Integrative	Strong influence	Strong influence	Ambivalence	Broad	Y	Y
Entrepreneurial discovery	Strong influence	Strong influence	Striving to improve	Narrow: dominated by OPEN	Y	Y
Innovation accomplishment	Strong influence	Weak influence	Striving to improve		n/a	Y
				Narrow: dominated by RIC	Y	n/a
Weakly coupled	Weak influence	Strong influence	Striving to improve	Weak overall context	n/a	N
Fragmented	Weak influence	Strong influence or Weak influence	Ambivalence or Striving to improve	Weak overall context	N	n/a

(Meijering et al., 2014). Often, there is a taken-for-grantedness of these indicators as facts (Power, 2021). However, by carefully considering their facticity, we recognize that metrics can serve as a transparency signal (Bermiss et al., 2014) and influence behaviour based on these metrics (Espeland & Sauder, 2007). As such, we suggest that they exert a strong (or weak) influence over stakeholders. For instance, from our study, high scores on the metric SMART (denoted as ‘fully-in’ in our analysis) indicated an unequivocal signal of capability that can be associated with performance. On the other hand, low scores on these conditions (denoted as ‘fully out’) indicated that these conditions were not a priority for stakeholders. Here, these results have a bearing on research that suggests that equivocal signalling may lead to decoupling (Espeland & Sauder, 2007), particularly because there are uncertain relations between the metrics and imprecise measures of sustainability. Therefore, we suggest that future research could further consider facticity, which may open a space for deeper reflection on the decoupling effects of commensurated frameworks and the complex power of metrics as transparency signals (Appio et al., 2019). However, we recognize that more work is needed to tighten the concept of facticity (Power, 2021).

Reflecting on the ‘reactivity to rankings’ dimension, research exemplifies concerns about positions in rankings that can affect competitive organizational behaviour. While our study found that striving for a higher rank, when coupled with strong openness and entrepreneurial support (SDG2), can lead to high performance (e.g. Vienna), our results also suggest that there may be a legitimacy façade (MacLean & Behnam, 2010), where attention to rankings would come at the expense of aligning activities for smart city development, leading to ambivalence to the metric (~SDG1) (Sauder & Espeland, 2009). However, our work also builds on calls to examine multiple facets of reactivity to rankings and study the different ways they affect organizational behaviour (Pollock et al., 2018). Moreover, the configurational nature of our study also allowed us to investigate the different reactions to rankings in settings with multiple metrics and how these effects unfold in

relation to performance. While it remains inconclusive, our work shows that reactivity to rankings does matter to performance in domains with multiple and different metrics.

Finally, regarding the dimension of context, as Mori and Christodoulou (2012) suggested, frameworks may wrongly imply that their dimensions are unrelated to each other and their context. However, contextual embedding is required to understand how configurations arise (Geels, 2002). Our findings confirm that what matters is the context that shapes the specific relationship between different conditions' effects on performance (Schot & Geels, 2008). We show that context can vary from being narrow in scope, typically focusing on a specific contextual condition, such as the QoG, to being broad in scope, also including the regional innovation system, highlighting context in a more holistic way. We also show that city conditions and context are simultaneously affected and connected. This emphasizes the value of a multi-level perspective (Mora et al., 2020). Thus, we advance an understanding of frameworking for smart city performance that is multi-level and complex (Schot & Geels, 2008). Further research could explore the nestedness of levels and the inter-connections of context and conditions to further understand how frameworks both enable and constrain organizational action. For instance, asymmetry identified in the findings shows that the outcome measure (SDG, QoL) cannot be fully inferred from their constitutive parts analysed in isolation or without context. Insofar as the outcomes (SDG, QoL) are highly prized by stakeholders, revealing that configurations of conditions and context characterize high (or low) performance have the potential to uncover dynamics of the marketization of sustainability performance (Callon & Muniesa, 2005).

Overall, our study makes a modest contribution to scholarship on commensuration. Organizational research on commensuration has focused mostly on reactivity to metrics and how reactivity often results in uneven performance (Espeland & Sauder, 2007), with some beginning to address commensuration in terms of facticity (Power, 2021). However, our approach attempts to capture both facticity and reactivity, highlighting that metrics are subject to different modes of reflection (Bermiss et al., 2014). Indeed, our configurational approach helped to theorize the notion of frameworking as a process of commensuration that not only produces metrics that are transparent but is also formative of organizational stakeholders' reaction to the frameworks and rankings produced. Therefore, future research on commensuration could focus on approaches that aim to strike a balance between using metrics as transparency signals of capability (Carlos & Lewis, 2018) and reactivity to rankings (Espeland & Stevens, 1998) that seeks to influence organizational behaviour (Mora et al., 2020). Further work could be undertaken on the concept of frameworking as a particular commensuration phenomenon, which is not just particular to smart cities. However, we also observed that an understanding of frameworking is hindered if studies disconnect reflections between transparency and reactivity. Therefore, we suggest that further studies examine theorizing that can embrace multiple reflective positions. Our specific approach allowed us to focus on the different reflections as configurations of encounters with the metrics. We acknowledge existing work in this area (cf. Bermiss et al., 2014; Orlikowski & Scott, 2014).

In terms of practical implications, the ideas and topics related to enumerating frameworks are not simply academic concerns but are of high relevance in practice as well. For instance, extant research shows that commensuration, including of frameworking, is hindered by a lack of 'input legitimacy' (Mena & Palazzo, 2012). Here, scholars have consistently noted that enumerating metrics and frameworks rarely involves broader stakeholder engagement. Broadening engagement may provoke deeper interrogations of incommensurability and transparency of the metrics that inform frameworks (Espeland & Stevens, 1998), challenging the sense of taken-for-grantedness of indicators as accurate representations of complex capabilities. Extending engagement might also suggest that reaction to performance in the rankings is not a simple choice but that stakeholders should embrace the opportunities that rankings provide. Nevertheless, our study shows that

reactivity to rankings alone will not produce satisfactory outcomes. Benefits will arise if reactivity is linked to well-thought-through interventions and governance.

In terms of smart city strategizing, indices and rankings are likely to rouse a range of nuanced and different strategic reactions. For instance, stakeholders may use them to sharpen a city's specific profile and define strategies for sustainable development (Acuto et al., 2021). Such use is from a predominantly functionalistic point of view. Here, frameworks are assumed to reveal their explanatory power and applicability to smart city strategic planning. However, frameworks also induce reactivity. This raises questions about whether frameworks are a valuable instrument for smart city stakeholders. For instance, stakeholders may ignore them if their city is poorly ranked (Giffinger et al., 2010). They may be adopted and abided by 'in good faith' or with superficial compliance, without scrutiny of the link between the metrics advocated and the strategies undertaken. These actions are also challenging, particularly when witnessing a proliferation of frameworks that bring together multiple metrics. In other words, careful attention needs to be paid to the normative questions regarding the objectivity and transparency aspects of smart city metrics and the different underlying reasons and motivations for reactivity to indices and rankings. Our study implies that frameworking as commensuration uncovered through a configurational approach will require stakeholders to employ a context-sensitive outlook in their attempts at strategizing. This should enable city stakeholders to find their position within ongoing smart city competition and initiate a learning process that makes their strategies more contextually oriented. For instance, this is important given our finding that the presence of openness, tolerance and trust is associated with both high and low performance towards the SDGs.

Finally, our study has some limitations. First, our sample was restricted to EU countries and therefore subject to prevailing institutional pressures. Urban sustainability is not confined to the EU but is a universal challenge set out by the United Nations (2020). However, we believe that our study is valuable for other regions, and we suggest further research in other regional contexts. Another limitation of our study is whether the causal conditions that apply to large cities are appropriate for cities of different sizes. Prior research suggests that city size matters (Giffinger et al., 2010) in how smart city development and sustainability unfolds, while other studies caution against attributing too much importance to size alone (Capello & Camagni, 2000). Therefore, there is some equivocality about whether the size of cities matters to smart city performance. We believe future studies should further consider the effects of city size and its implications for smart city development and sustainability to ensure that theoretical work specific to the development dynamics of smaller- and medium-sized cities can be addressed. Furthermore, we acknowledge that there might be methodological limitations that could provide fruitful avenues for further research. Future work should focus more closely on rankings and commensuration to get closer to how reputation affects smart city performance. For instance, fsQCA is often seen as only able to deal with cross-sectional analysis and poor for more longitudinal studies (Misangyi et al., 2016). Studies on smart cities would be enhanced with more longitudinal attention (Mora et al., 2020) to help shed important light on the continuing interest in urban sustainability.

Acknowledgements

We would like to thank the anonymous reviewers and the guest editors for their guidance during the review process.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Katharina Burger  <https://orcid.org/0000-0002-4348-5381>

Notes

- 1 Regional Competitiveness Index, Smart City Index 2020, European Cultural and Creative Cities Monitor, Easypark Smart City Index, European Digital City Index, IESE CIMI (Cities in Motion) Index, SDG Index and Dashboards Report for European Cities, Numbeo Quality of Life Index, OECD Regional data.
- 2 Cities between 250,000 and 5,000,000 inhabitants (i.e. L, XL and XXL in the EU-OECD definition).

References

- Abbott, Andrew (1988). Transcending general linear reality. *Sociological Theory*, 6, 169–186.
- Acuto, Michele, Pejic, Daniel, & Briggs, Jessie (2021). Taking city rankings seriously: Engaging with benchmarking practices in global urbanism. *International Journal of Urban and Regional Research*, 45, 363–377.
- Appio, Francesco P., Lima, Marcos, & Paroutis, Sotirios (2019). Understanding smart cities: Innovation ecosystems, technological advancements, and societal challenges. *Technological Forecasting and Social Change*, 142, 1–14.
- Bermis, Sekou Y., Zajac, Edward J., & King, Brayden G. (2014). Under construction: How commensuration and management fashion affect corporate reputation rankings. *Organization Science*, 25, 591–608.
- Brandtner, Christof, Höllerer, Markus A., Meyer, Renate E., & Kornberger, Martin (2016). Enacting governance through strategy: A comparative study of governance configurations in Sydney and Vienna. *Urban Studies*, 54, 1075–1091.
- Brunsson, Nils, Rasche, Andreas, & Seidl, David (2012). The dynamics of standardization: Three perspectives on standards in organization studies. *Organization Studies*, 33, 613–632.
- Callon, Michel, & Muniesa, Fabian (2005). Peripheral vision: Economic markets as calculative collective devices. *Organization Studies*, 26, 1229–1250.
- Capello, Roberta, & Camagni, Roberto (2000). Beyond optimal city size: An evaluation of alternative urban growth patterns. *Urban Studies*, 37, 1479–1496.
- Caragliu, Andrea, & Del Bo, Chiara F. (2019). Smart innovative cities: The impact of smart city policies on urban innovation. *Technological Forecasting and Social Change*, 142, 373–383.
- Carlos, W. Chad, & Lewis, Ben W. (2018). Strategic silence: Withholding certification status as a hypocrisy avoidance tactic. *Administrative Science Quarterly*, 63, 130–169.
- Chatterji, Aaron K., & Toffel, Michael W. (2010). How firms respond to being rated. *Strategic Management Journal*, 31, 917–945.
- Clementino, Ester, & Perkins, Richard (2020). How do companies respond to environmental, social and governance (ESG) ratings? Evidence from Italy. *Journal of Business Ethics*, 171, 1–19.
- Czarniawska, Barbara (2002). *A tale of three cities: Or the glocalization of city management*. Oxford: Oxford University Press.
- De Guimarães, Julio C., Severo, Eliana A., Felix Júnior, Luiz A., Da Costa, Wényka P., & Salmoria, Fernanda T. (2020). Governance and quality of life in smart cities: Towards sustainable development goals. *Journal of Cleaner Production*, 253, 119926.
- Denk, Thomas, & Lehtinen, Sarah (2014). Contextual analyses with QCA-methods. *Quality & Quantity*, 48, 3475–3487.
- Doty, D. Harold, Glick, William H., & Huber, George P. (1993). Fit, equifinality, and organizational effectiveness: A test of two configurational theories. *Academy of Management Journal*, 36, 1196–1250.
- Elsbach, Kimberly D., & Kramer, Roderick M. (1996). Members' responses to organizational identity threats: Encountering and countering the business week rankings. *Administrative Science Quarterly*, 41, 442–476.
- Espeland, Wendy N., & Sauder, Michael (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, 113, 1–40.

- Espeland, Wendy N., & Stevens, Mitchell L. (1998). Commensuration as a social process. *Annual Review of Sociology*, 24, 313–343.
- Fiss, Peer C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, 54, 393–420.
- Fombrun, Charles, & Shanley, Mark (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal*, 33, 233–258.
- Geels, Frank (2002). Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case study. *Research Policy*, 31, 1257–1274.
- Gerdin, Jonas, & Englund, Hans (2019). Contesting commensuration: Public response tactics to performance evaluation of academia. *Accounting, Auditing & Accountability Journal*, 32, 1098–1116.
- Gianelle, Carlo, Guzzo, Fabrizio, & Mieszkowski, Krzysztof (2020). Smart specialisation: What gets lost in translation from concept to practice? *Regional Studies*, 54, 1377–1388.
- Giffinger, Rudolf, Haindlmaier, Gudrun, & Kramer, Hans (2010). The role of rankings in growing city competition. *Urban Research & Practice*, 3, 299–312.
- Graffin, Scott D., & Ward, Andrew J. (2010). Certifications and reputation: Determining the standard of desirability amidst uncertainty. *Organization Science*, 21, 331–346.
- Gunn, Andrew (2018). Metrics and methodologies for measuring teaching quality in higher education: Developing the Teaching Excellence Framework (TEF). *Educational Review*, 70, 129–148.
- Hollands, Robert G. (2015). Critical interventions into the corporate smart city. *Cambridge Journal of Regions Economy and Society*, 8, 61–77.
- Huault, Isabelle, & Rainelli-Weiss, H el ene (2011). A market for weather risk? Conflicting metrics, attempts at compromise, and limits to commensuration. *Organization Studies*, 32, 1395–1419.
- Ivaldi, Enrico, Penco, Lara, Isola, Gabriele, & Musso, Enrico (2020). Smart sustainable cities and the urban knowledge-based economy: A NUTS3 level analysis. *Social Indicators Research*, 150, 45–72.
- Jessop, Bob, & Sum, Ngai-Ling (2000). An entrepreneurial city in action: Hong Kong's emerging strategies in and for (inter)urban competition. *Urban Studies*, 37, 2287–2313.
- Katz, Daniel, & Kahn, Robert L. (1966). *The social psychology of organizations*. New York: John Wiley & Sons.
- Kim, Jerry W. (2020). Halos and egos: Rankings and interspecialty deference in multispecialty US hospitals. *Management Science*, 66, 2248–2268.
- Kornberger, Martin, & Carter, Chris (2010). Manufacturing competition: How accounting practices shape strategy making in cities. *Accounting, Auditing & Accountability Journal*, 23, 325–349.
- Levin, Peter, & Espeland, Wendy N. (2002). Pollution futures: Commensuration, commodification, and the market for air. In Andrew Hoffman & Marc Ventresca (Eds.), *Organizations, policy and the natural environment: Institutional and strategic perspectives* (pp. 119–148). Stanford, CA: Stanford University Press.
- Leydesdorff, Loet, & Deakin, Mark (2011). The triple-helix model of smart cities: A neo-evolutionary perspective. *Journal of Urban Technology*, 18(2), 53–63.
- MacLean, Tammy L., & Behnam, Michael (2010). The dangers of decoupling: The relationship between compliance programs, legitimacy perceptions, and institutionalized misconduct. *Academy of Management Journal*, 53, 1499–1520.
- Marginson, Simon, & Van der Wende, Marijk (2007). To rank or to be ranked: The impact of global rankings in higher education. *Journal of Studies in International Education*, 11, 306–329.
- Mazmanian, Melissa, & Beckman, Christine M. (2018). “Making” your numbers: Engendering organizational control through a ritual of quantification. *Organization Science*, 29, 357–379.
- Meijer, Albert (2018). Datapolis: A public governance perspective on “smart cities”. *Perspectives on Public Management and Governance*, 1, 195–206.
- Meijering, Jurian V., Kern, Kristine, & Tobi, Hilde (2014). Identifying the methodological characteristics of European green city rankings. *Ecological Indicators*, 43, 132–142.
- Mena, S ebastien, & Palazzo, Guido (2012). Input and output legitimacy of multi-stakeholder initiatives. *Business Ethics Quarterly*, 22, 527–556.

- Mennicken, Andrea, & Espeland, Wendy N. (2019). What's new with numbers? Sociological approaches to the study of quantification. *Annual Review of Sociology*, 45, 223–245.
- Meyer, Alan D., Tsui, Anne S., & Hinings, Christopher R. (1993). Configurational approaches to organizational analysis. *Academy of Management Journal*, 36, 1175–1195.
- Misangyi, Vilmos F., Greckhamer, Thomas, Furnari, Santi, Fiss, Peer C., Crilly, Donal, & Aguilera, Ruth (2016). Embracing causal complexity: The emergence of a neo-configurational perspective. *Journal of Management*, 43, 255–282.
- Mora, Luca, Bolici, Roberto, & Deakin, Mark (2017). The first two decades of smart-city research: A bibliometric analysis. *Journal of Urban Technology*, 24(1), 3–27.
- Mora, Luca, & Deakin, Mark (2019). *Untangling smart cities: From utopian dreams to innovation systems for a technology-enabled urban sustainability*. Amsterdam: Elsevier.
- Mora, Luca, Deakin, Mark, Zhang, Xiaoling, Batty, Michael, de Jong, Martin, Santi, Paolo, & Appio, Francesco P. (2020). Assembling sustainable smart city transitions: An interdisciplinary theoretical perspective. *Journal of Urban Technology*, 28, 1–27.
- Mori, Koichiro, & Christodoulou, Aris (2012). Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI). *Environmental Impact Assessment Review*, 32(1), 94–106.
- Orlikowski, Wanda J., & Scott, Susan V. (2014). What happens when evaluation goes online? Exploring apparatuses of valuation in the travel sector. *Organization Science*, 25, 868–891.
- Paroutis, Sotirios, Bennett, Mark, & Heracleous, Loizos (2014). A strategic view on smart city technology: The case of IBM smarter cities during a recession. *Technological Forecasting and Social Change*, 89, 262–272.
- Patorniti, Nicholas P., Stevens, Nicholas J., & Salmon, Paul M. (2018). A sociotechnical systems approach to understand complex urban systems: A global transdisciplinary perspective. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 28, 281–296.
- Pollock, Neil, D'Adderio, Luciana, Williams, Robin, & Leforestier, Ludovic (2018). Conforming or transforming? How organizations respond to multiple rankings. *Accounting, Organizations and Society*, 64, 55–68.
- Porac, Joseph F., & Thomas, Howard (1990). Taxonomic mental models in competitor definition. *Academy of Management Review*, 15, 224–240.
- Power, Michael (2021). Modelling the micro-foundations of the audit society: Organizations and the logic of the audit trail. *Academy of Management Review*, 46, 6–32.
- Ragin, Charles C. (2000). *Fuzzy-set social science*. Chicago, IL: University of Chicago Press.
- Ragin, Charles C., Drass, Kriss A., & Davey, Sean (2006). *Fuzzy-set/qualitative comparative analysis 2.0*. Tucson, AZ: Department of Sociology, University of Arizona.
- Rao, Hayagreeva (1994). The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895–1912. *Strategic Management Journal*, 15, 29–44.
- Richter, Chris, Kraus, Sascha, & Syrjä, Pasi (2015). The smart city as an opportunity for entrepreneurship. *International Journal of Entrepreneurial Venturing*, 7, 211–226.
- Roberts, John (2018). Managing only with transparency: The strategic functions of ignorance. *Critical Perspectives on Accounting*, 55, 53–60.
- Rowley, Timothy J., Shipilov, Andrew V., & Greve, Henrich R. (2017). Board reform versus profits: The impact of ratings on the adoption of governance practices. *Strategic Management Journal*, 38, 815–833.
- Sauder, Michael, & Espeland, Wendy N. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review*, 74, 63–82.
- Schneider, Anselm, Wickert, Christopher, & Martí, Emilio (2017). Reducing complexity by creating complexity: A systems theory perspective on how organizations respond to their environments. *Journal of Management Studies*, 54, 182–208.
- Schneider, Carsten, & Wagemann, Claudius (2012). *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*. Cambridge: Cambridge University Press.
- Schot, Johan, & Geels, Frank (2008). Strategic niche management and sustainable innovation journeys: Theory, findings, research agenda, and policy. *Technology Analysis & Strategic Management*, 20, 537–554.

- Schultz, Majken, Mouritsen, Jan, & Gabrielsen, Gorm (2001). Sticky reputation: Analyzing a ranking system. *Corporate Reputation Review*, 4, 24–41.
- Sengers, Frans, Wieczorek, Anna J., & Raven, Rob (2019). Experimenting for sustainability transitions: A systematic literature review. *Technological Forecasting and Social Change*, 145, 153–164.
- Sharifi, Ayyoob (2020). A typology of smart city assessment tools and indicator sets. *Sustainable Cities and Society*, 53, 101936.
- Sharkey, Amanda J., & Bromley, Patricia (2015). Can ratings have indirect effects? Evidence from the organizational response to peers' environmental ratings. *American Sociological Review*, 80, 63–91.
- Simmons, Geoff, Giraldo, Jorge E., Truong, Yann, & Palmer, Mark (2018). Uncovering the link between governance as an innovation process and socio-economic regime transition in cities. *Research Policy*, 47, 241–251.
- Slager, Rieneke, & Gond, Jean-Pascal (2022). The politics of reactivity: Ambivalence in corporate responses to corporate social responsibility ratings. *Organization Studies*, 43, 59–80.
- Slager, Rieneke, Gond, Jean-Pascal, & Crilly, Donal (2021). Reactivity to sustainability metrics: A configurational study of motivation and capacity. *Business Ethics Quarterly*, 31, 275–307.
- United Nations (2020). *Voluntary Local Reviews*. Department of Economic and Social Affairs; Sustainable Development. <https://sdgs.un.org/topics/voluntary-local-reviews>
- White, Leroy, Lockett, Andy, Currie, Graeme, & Hayton, James (2021). Hybrid context, management practices and organizational performance: A configurational approach. *Journal of Management Studies*, 58, 718–748.
- Wijen, Frank (2014). Means versus ends in opaque institutional fields: Trading off compliance and achievement in sustainability standard adoption. *Academy of Management Review*, 39, 302–323.
- Yigitcanlar, Tan, Han, Hoon, Kamruzzaman, Md, Ioppolo, Giuseppe, & Sabatini-Marques, Jamile (2019). The making of smart cities: Are Songdo, Masdar, Amsterdam, San Francisco and Brisbane the best we could build? *Land Use Policy*, 88, 104187.

Author biographies

Leroy White is professor of management at the University of Exeter Business School. His research interests are in management practices, leadership, and partnership working. Empirically, he works on public and not-for-profit organizations.

Katharina Burger is a senior lecturer in management at the University of Bristol, UK. Her research focuses on multi-stakeholder decision support processes, particularly related to city futures.