Replicate and generalize to make urban research coherent

Citation for published version:

Digital Object Identifier (DOI):
10.1080/12265934.2024.2382706

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
International Journal of Urban Sciences

Publisher Rights Statement:
© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Replicate and generalize to make urban research coherent

A. Sudmant, F. Creutzig & Z. Mi

To cite this article: A. Sudmant, F. Creutzig & Z. Mi (07 Aug 2024): Replicate and generalize to make urban research coherent, International Journal of Urban Sciences, DOI: 10.1080/12265934.2024.2382706

To link to this article: https://doi.org/10.1080/12265934.2024.2382706

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 07 Aug 2024.

Submit your article to this journal

Article views: 21

View related articles

View Crossmark data
Replicate and generalize to make urban research coherent

A. Sudmant a,b, F. Creutzig c and Z. Mi d

a University of Edinburgh, Edinburgh Climate Change Institute, Edinburgh, UK; b Earth and Environment Department, University of Leeds, Leeds, UK; c Mercator Research Institute on Global Commons and Climate Change, Sustainability Economics of Human Settlements at Technical University Berlin, Berlin, Germany; d The Bartlett School of Sustainable Construction, University College London, London, UK

ABSTRACT
Scientific approaches to urban research are a century in the making but have long remained tangled in debates and disagreements. Here, we reflect on how current enthusiasm for a scientific urbanism resurfaces old conflicts and suggest that a diverse set of arguments revolve around two concepts: replication and generalizability. We argue that addressing these challenges directly could play an important role in advancing empirical research. New empirical methods, including those coming from artificial intelligence, offer means of radically accelerating urban research, but also risk a revisiting of longstanding disagreements. Applying replication and generalizability as a conceptual bridge may be key to some of these new methods contributing to a more integrated and interdisciplinary urbanism. Beyond empirical urbanism, replication and generalizability may enable the development of a more integrated, systematic urban field by supporting the translation of methodologies, findings, and theories between disciplines. We conclude by cautioning that further work is needed to understand how replication and generalizability apply across the urban field (significant areas of which are not addressed in this paper) and by calling for authors to test the value of replication and generalizability in their areas of urban expertise.

ARTICLE HISTORY
Received 8 August 2023
Accepted 28 June 2024

KEYWORDS
Cities; urban areas; urban science; science of cities; city science; interdisciplinarity; replication; generalizability

Highlights

– Scientific urbanism is failing to move past longstanding conflicts and contradictions.
– The boundary objects replication and generalisability can support a more coherent empirical urbanism.
– Replication engages with the extent urban phenomena can be captured by empirical analyses.
– Generalisability engages with the extent a studies’ insights are useful in other contexts.
– Further work is needed to understand the roles of replication and generalisability in the wider urban field.

CONTACT
A. Sudmant a.sudmant@ed.ac.uk

Supplemental data for this article can be accessed online at https://doi.org/10.1080/12265934.2024.2382706.
1. Introduction

An organism, a tree, a house, a work of art, a machine, a political entity: the metaphors we use and have used for millennia to describe cities capture the complexity and the enduring puzzle they pose for our understanding.

In the last 25 years or so rejuvenated efforts to bring empirical methods to urban research, including the development of a so-called Urban Science, have provided a new set of tools and approaches to urban questions (Bai et al., 2017; Batty, 2021; L. M. A. Bettencourt & Zünd, 2020; Creutzig, 2015; Hsu et al., 2020; Ivanova & Eichner, 2021). Behind a rise in the application of these approaches are advances in analytical and statistical methods (Gelman & Vehtari, 2021), and ever more available and detailed urban data (Batty, 2020; Creutzig et al., 2019; Sethi & Creutzig, 2023; Sudmant et al., 2022). An increase in focus on cities from academics of diverse backgrounds is also the result of the rise of the political, demographic, cultural and environmental importance of cities (Bai et al., 2018).

Unique to this latest turn towards a quantitative urbanism is its focus on providing practical insight into the challenges humanity faces in the twenty-first century, including the Sustainable Development Goals and Climate Change. Creutzig (2015), for example, uses a dataset of urban energy use and machine learning approaches to reveal wedges of potential GHG emissions reductions in cities. Cheng, Mi, Sudmant, and Coffman (2022) reveal the relationship between emissions from urban areas in China and the urban population after accounting for economic growth, economic structure and geography.

Enthusiasm for a scientific urbanism is therefore understandably high. A 'Science of Cities' is seen to have the potential to help us respond to a global Climate Emergency (Creutzig et al., 2019) and to support the wider project of a transition to sustainability (Acuto, Parnell, & Seto, 2018; Purvis, 2021). It is thought to help us to understand the combined technical, social, political and ecological nature of cities (Folke, Biggs, Norström, Reyers, & Rockström, 2016; Grimm, Cook, Hale, & Iwaniec, 2015; McPhearson et al., 2016; Meerow, Newell, & Stults, 2016) and to account for the interactive and dynamic complexity of urban form (Barth, 2005). It is also seen as a foundation for understanding the exchange of resources across multiple scales and the flows of goods, people, information and ideas as they can be generalized across city size and scale (Batty, 2012a, 2019).

Empirically led approaches to the urban, however, face today the same challenges they have faced across the twentieth century. Datasets, though vastly superior to those of the early century, nearly always miss aspects of the urban, and key urban ideas and concepts can be challenging to capture and test by empirical means. In addition, cities, both in the real and academic worlds, are a meeting point. The richness of the city, as well as its complexity, comes from the range and diversity of things it encompasses, demanding means of interdisciplinary collaboration and learning for our understanding to progress.

In this paper, we suggest that the concepts of replication and generalizability are key to advancing empirical urban research. The concept of replication engages with the extent to which urban phenomena can be captured by empirical analyses. Central to this challenge is the qualitative nature of many urban concepts, such as ‘capacity’ and ‘governance’. Generalizability, in contrast, explores the extent to which an insight from a study is useful in a wider context. Missing information, including aspects of the urban not in datasets and urban areas not in datasets, challenges generalizing from research.
Specifically, we make the case that empirical urban research can be made more relevant and impactful when the questions posed and results produced directly reference the challenges of replication and generalizability. Thinking beyond empirical research, we also suggest that the concepts of replication and generalizability may contribute more generally to advancing a more systematic, coherent and interdisciplinary urbanism by serving as boundary objects between epistemic communities (Akkerman & Bakker, 2011; Star & Griesemer, 1989).

We start in Sections 2 and 3 by exploring existing urban literature and suggesting that longstanding debates can be framed by replication and generalizability. This review of literature, focusing on fields and authors who have brought empirical methods to urban research, is necessarily short and leaves out many key contributions and debates. Following this, Sections 5 and 6 explore replication and generalizability in more detail by learning from authors in the fields of sociology, data science and psychology where replication and generalizability have been conceptually developed. In Section 7, we present a framework built around replication and generalizability to ground the concepts in urban questions and explore where replication and generalizability are independent and exist together. Section 8.1 explores how replication and generalizability can help to advance empirical urban research. Section 8.2 considers the role of replication and generalizability, as well as boundary objects more generally, for advancing a wider urban field. Section 9 provides conclusions.

Our analysis maintains some bias in the examples and references towards urban environmental research, reflecting the expertise of the authors. Further analysis is needed to understand how replication and generalizability play a role in wider fields of urban research, including urban metabolism (c.f Inostroza & Zepp, 2021), morphology (c.f Artmann, Inostroza, & Fan, 2019; Lemoine-Rodríguez, Inostroza, & Zepp, 2020), sociology (c.f Sassen, 2018), planning and design (c.f Plater-Zyberk & Lombard, 2010), governance (c.f Bulkeley, 2010) and economics (c.f Ala-Mantila, Heinonen, & Junnila, 2014; Colenbrander et al., 2017), among other fields.

2. Urban science in the twentieth century: discovery and re-discovery of the city

Systems of urban planning, engineering, architecture and governance have existed for millennia. The cities of Mohenjo-daro and Harappa in the Indus Valley (circa 2600–1900 BCE) had advanced municipal water management systems that indicate advanced engineering capacities and systems of governance (Jansen, 1985). Xianyang (circa 250 BCE) was built with integrated water systems, roads, and marketplaces, offering an early example of integrated urban planning (Zhang & Wang, 2019). Timbuktu, under the reign of Mansa Musa in the fourteenth century, was a scholarly and architectural hub, with the Djinguereber Mosque being a notable example of urban architecture that combined local traditions with Islamic influences (Pradines, 2022).

An urbanism that bridges academic disciplines, however, is a more recent ambition. ‘Discovery’ of the city by a diverse set of academic fields, and invention and reinvention of an ‘urban science’, occurred periodically throughout the twentieth century.

Drawing on key concepts and theories from the biological sciences, sociologists at the University of Chicago, including Robert Park, Ernest Burgess and Roderick McKenzie
seeded urban ecology in the 1920s (Wu, 2014). Defined initially as ‘the study of the relationship between people and their urban environment’ (Park, 1915, p. 578), urban ecologists applied concepts from ecology and sociology, including dominance, succession and competition, to the study of cities (Hollingshead, 1940; Park, 1915).

Successors of this field today view cities themselves as ecosystems (Grimm, Grove, Pickett, & Redman, 2000; McPhearson et al., 2016). As key actors in this ecosystem, human lifestyles and living patterns, and the social structures and political processes human life gives rise to, are therefore key aspects of analysis. Key concepts and theories that provide the foundations for this research include socioecological resilience, ecosystem services, complex adaptive systems, and social-ecological systems (Wu, 2014). Urban ecology as a discipline has therefore dramatically expanded its scope of study in response to an understanding of the multiple and interconnected dimensions of urban ecosystems (Pickett et al., 2013).

Independent of the work of urban ecologists, advances in computational sciences and an ‘avalanche of printed numbers’ (Hacking, 1982), drew practitioners from physics, mathematics and economics to urban questions in the years after WWII (Alberti, 2017). In most cases, these authors treated cities as organized from the top down and independent from their wider environment (Chadwick & Chadwick, 1971). As this work developed, ‘urban modelling’ as mathematical or statistical representations of an urban system came to replace ‘urban modelling’ as an architectural representation of the city (Batty & Torrens, 2001).

In the decades that followed, urban quantitative geography emerged as a distinct sub-field of urban geography. Quantitative geography today seeks a more complex understanding of the urban than the top-down approaches of the 1950s and 1960s, centring on the identifiable and recurring socioeconomic patterns that emerge where people agglomerate. Importantly, these patterns connect urban quantities, in terms of population and infrastructures, to urban qualities, including the socio-economic character of urban areas (Carneiro, 2000), and a wide range of social, political, behavioural and technological aspects of the urban (Batty, 2020; Cottineau, Hatna, Arcaute, & Batty, 2017; Heppenstall, Crooks, See, & Batty, 2011).

While the origin of urban-focused economics and physics is often traced to the development of ‘quantitative geographies’ in the 1950s and 1960s, or in some cases even earlier (Ball, 2002), distinct areas of research developed only more recently. In economics, Berry (1999) describes economists’ ‘discovery’ of the ‘puzzling regularity of the urban hierarchy’ (Krugman, 1991, p. 399) in the early 1990s, part of a broader movement that was termed ‘New Economic Geography’.

Led by developments in complexity sciences, physicists’ focus on the city emerged in the early 2000s (Batty, 2012b). Physicists have followed statistical geographers and economists in the search for ‘coarse grained regularities’, ‘universals’, and ‘laws’ of the urban (L. M. Bettencourt, Lobo, & Strumsky, 2007; L. Bettencourt & West, 2010; Liotta, Vigué, & Creutzig, 2023; Louf & Barthelemy, 2015; Pumain, 2018; Vacchiani-Marcuzzo, 2005). Statistical geographers refer to ‘many sciences of cities’ (Batty, 2020, p. 16) and New Urban Geography scholars consider economies at different stages of development, urban physicists, however, seek more fundamental characterizations of the urban that apply universally. Lobo et al. (2020, p. 10), for example, describe a science of cities as
[capturing] fundamental processes that lie at the core of all human spatial agglomerations, whether these are past or present, agrarian or industrial, in developed economies or developing countries … In the same way that evolution applies to the fossil record as well as living populations.

3. Methodological progress, epistemological regress: axes of the debate

The preceding paragraphs provide only a brief outline of disciplinary exchange between urban fields of research and disciplines that have not traditionally focused on urban settings; more detailed discussion can be found in Ball (2002), Sennett (1996), Purvis (2021) and Barnes (2013). This introduction allows us, however, to make two high-level observations.

First, there has been a longstanding and well-documented project to bring approaches and theories from the natural sciences to urban questions. These exchanges have been encouraged by empirical breakthroughs, including those following World War II and recent advances in complexity and statistical sciences, but date back at least as far as the origins of statistical physics in the nineteenth century (see Ball, 2002). These exchanges have come from disciplines that traditionally have not made cities a distinct area of research, such as ecology, economics, and physics, but also from within urban studies. For example, urban theorists including Geddes, Mumford, Park, and Jacobs drew on Darwin’s theory of evolution to describe cities’ growth and change (Sennett, 1996). These contributions were often marked by ‘an exclusionary discourse to mark the new from the old’ (Derudder & van Meeteren, 2019, p. 6), and the mantle of ‘scientific’ has been used both to champion and diminish new work.

Second, frictions between established and novel urban research have had two common themes. The first of these focuses on the extent to which novel, and especially quantitative, methodologies, include key aspects of urban systems. These can be seen in the criticism directed at systems theorists of the 1950s and 1960s that cities could not be studied as discrete enclosed systems (Batty, 2012b) and in criticism in the 1970s that non-numerical context is left out in quantitative studies of the urban (Ley & Samuels, 2014). Healey (1987) was among the authors in the 1980s emphasizing that quantitative urbanism ignored underlying social and economic processes while Brenner and Schmid (2015) and Gleeson (2014) are among contemporary authors critical of Big Data urbanism and its role as ‘part of a broader technoscientific ideology that aims to depoliticize urban life’ (Gleeson, 2014, p. 348).

Another theme of criticism has focused not on the information that is missing from analyses, but on how we understand the information that is included. Key authors who have focused on this issue include Sayer (1984), Harvey (1992) and Olsson (1980). Sayer (1984), addressing quantitative social science research from a critical realist perspective, emphasizes that numbers are always marked by prior theorization, while Olsson (1980) argues that data necessarily represent existing social and institutional arrangements, mobilizations of power, and political agendas. These criticisms are brought forward in contemporary work by Barnes (2013) in the context of ‘Big Data’, who argues these challenges lead the field to be inherently conservative as it is constrained by the world as it is rather than what it could be, or should be.

Importantly, the parties in these disagreements cannot be separated into so-called empirical positivists and ‘those trained to see subjectivities’ (Duminy & Parnell, 2020,
p. 4), which is to say the clichéd division between natural and social scientists. Social scientists, for example, have been divided over ‘universalizing urbanism’ (Robinson & Roy, 2016; Storper & Scott, 2016; Robinson & Roy, 2016) and the extent to which GIS data can recapitulate power, political agendas and existing societal structures (Schuurman, 2000; Schuurman & Pratt, 2002).

4. Defining replication and generalizability

The key themes of disagreement within and between urban scholars and other disciplines offer a means of defining the challenges the field faces as the body of empirical research grows.

The replication problem engages with the extent to which urban phenomena can be captured by empirical analyses. Central to this challenge is the qualitative nature of many urban concepts and the imperfect means of operationalizing these concepts that researchers are forced to employ. Urban capacity, civic pride, and the quality of urban governance are examples of concepts for which urban researchers are forced to develop proxies in order to conduct empirical analyses (c.f Chan & Hu, 2004; Collins, 2016; Onishi, 1994).

The generalizability problem, engages with the challenge of identifying insights with broad-based relevance. Missing information, including unquantified aspects of cities as well as cities that are not quantified in datasets, creates sharp boundaries around the kind of broad-based conclusions analyses can make. The diversity of contexts and settings that we would collectively identify as urban, the continuous change that characterizes urban areas and the complexity of urban systems raise challenges for the identification of generalizable ‘rules’ of cities that move from being methodological to epistemological. To the extent that findings across groups of urban areas can be found, to what extent are they dependent on the way we define the urban?

Replication and generalizability challenges are therefore closely connected, and in some cases overlap. Differentiating between these challenges, however, is critical to the advancement of empirical urban research.

5. The generalizability crisis: simple rules, complex cities, and the need for defining boundaries of analysis

Science is famously the art of ‘discerning what we may with advantage omit’ (Popper, 1992, p. 44). To this end Lobo et al describe a Science of Cities

[capturing] fundamental processes that lie at the core of all human spatial agglomerations, whether these are past or present, agrarian or industrial, in developed economies or developing countries … In the same way that evolution applies to the fossil record as well as living populations. (2020, p. 10)

But dispensing with those aspects that are of a place or a time requires capturing these aspects as information in the first place.

Content validity considers the degree to which an analysis captures a subject or phenomenon in detail sufficient for analysis to draw conclusions. Missing aspects of a subject or phenomena, or bias in these are key barriers to content validity and a key
Subjects or phenomena that present in different ways in different contexts will therefore be the most likely to lead to generalizability problems. Generalizability errors will arise as a matter of degree: until there is both a well-defined urban science and a well-established definition of what is and what is not part of the urban, tasks unlikely to be resolved in the near future.

In an urban context, issues with content validity arise in two key ways: via limitations in observations of a variable and through missing variables. The first of these arises where datasets contain bias. Some Science of Cities literature (e.g. Lobo et al., 2020) claim empirical datasets and published research bias was mainly an issue of the nineteenth and twentieth centuries. In conflict with this claim, wider literature documents how information on the urban continues to be focused towards the Global North (Nagendra, Bai, Brondizio, & Lwasa, 2018), towards male perspectives and experiences (Khreis et al., 2016), towards the economic and financial over the social and environmental (Gouldson, Sudmant, Khreis, & Papargyropoulou, 2018), and towards those with higher socio-economic status (Colenbrander et al., 2017).

Biases in datasets can also be specific to their origins, owners and creators. While services that rely on Big Data, including mapping services, social media algorithms and streaming music, are used by billions, the scientific value of the data on which they rely does not necessarily follow from their social utility. The collecting, cleaning and sorting of data can alter the understanding of phenomena in ways that researchers, and in many cases data providers, are unaware of (Zou & Schiebinger, 2018). Big Data often contain a restricted number of variables, therefore requiring inferences with explanatory characteristics that can simultaneously compound the impact of, and obscure the source of, biases (Hu & Jin, 2017). And what data is available, under what circumstances and for what parties can affect the questions being asked, and by extension the progress of a science that follows (Albino, Berardi, & Dangelico, 2015; Docherty, Marsden, & Anable, 2018; N. Wang & Ma, 2021).

A second source of content validity error arises as a consequence of missing aspects of the urban. The number of features included in an analysis is determined by the availability of data and the theoretical framing of the question being addressed by the author. Underappreciated, it can also arise when researchers reduce the number of variables in their analysis as a means of addressing variable interaction and collinearity (Tosh et al., 2021).

The degree to which a reduction in the number of explanatory features is necessarily a ‘feature’ or ‘bug’ depends on your perspective on urban questions. Gelman (2018) raises the question of whether controlling for social factors or ‘bracketing the social’ leads to a deeper understanding of phenomena for which there is a fundamental social component, or whether it leads to the ‘evaporation’ of those variables from the analysis (p. 274). A ‘reductive tendency’, found in research on climatology, engineering, and physics, among other fields, however, has been found to produce misleading and sometimes dangerous results (Gras, Conger, Jenkins, & Gras, 2019).

Purvis (2021) notes that so-called neutral approaches to urban science are key to enabling the development of systematic approaches. When the tools of natural sciences are applied to the social sciences, Yarkoni and Westfall (2017) note that there is often an implicit notion that causes and effects are so intertwined that their separation is more a philosophical than practical exercise. Whether or not a neutral or ‘de-linked’ analysis of a
phenomenon with social aspects is possible, however, is questioned by authors who posit that political and normative considerations are simply embedded rather than removed in scientific methods and arguments (Gelman, 2019b; Grätz, 2018; Hardwicke & Ioannidis, 2019). Furthermore, the analysis focused on the proximate aspects of the urban and framed towards understanding and engaging with poorly defined topics such as ‘urban transformations’ may be fertile ground for being co-opted by established discourses and problematizations that are opposed to structural change (Westman & Castán Broto, 2022).

Two potential causes of generalizability error can therefore arise. First, where aspects of the urban are controlled for, or ‘bracketed’, proximate factors can be misinterpreted as causal drivers of phenomena. In other words, patterns can be construed with processes. Second, where data is absent of or invariant in aspects of the urban – social, political, or cultural or otherwise – analysis can err by not clearly defining the boundaries of analysis. In these cases, there is a critical need for literature to more clearly define the subjects and questions being considered.

In addition, there are two ways generalizability errors can emerge. In the first case, the authors themselves fail to unpick causal and correlated factors, or to identify the boundaries of their analysis. In the second case, the citing and reciting of research can lead the wider research community and the general public to bend and stretch the conclusions of research.

6. The replication crisis: can we model the essence of the urban?

Over the last two decades, concerns have been raised in a number of fields around the extent to which research findings, including longstanding and sometimes foundational research, can be replicated by new studies. The majority of this discussion has focused on shortfalls in research practices, and around selection pressures related to funding, publication and academic advancement (Gelman, 2019b; Shrout & Rodgers, 2018; Yarkoni, 2022).

In the vast majority of these cases, ‘un-replicable studies’ are the result of honest mistakes and misjudgements. A wider form of replication challenge can also occur, however, due to an overarching theoretical challenge: Many of the key concepts we draw on to describe and understand urban areas have no obvious empirical form.

In the following, we differentiate between two kinds of replication errors. The first kind of error is made due to outright mistakes on the part of researchers. These mistakes can be methodological, computational, or conceptual. The second kind of replication error emerges due to the challenge of operationalizing aspects of the urban, both those that makeup the dependent and independent elements of analytical research.

6.1. Procedural flexibility and the garden of forking paths: would a small change affect our conclusions?

Choices made around approaches to data collection, analysis and presentation have implications for the findings of research. This tautological statement is widely understood. Simultaneously, however, the freedom of researchers to explore sources of information, datasets, methods, what Simmons describes as ‘researcher degrees of freedom’
(2011), is widely celebrated and typically deemed a necessary part of the research process. Problematically, very few of the choices researchers make are documented, much less are choices considered in a systemic way. Consequently, research frequently progresses down a ‘garden of forking paths’ where it becomes impossible for a researcher to determine if choices led to minor or massive implications (Gelman & Loken, 2014).

Research has found that the consequences of seemingly minor changes in research processes can lead to substantial changes in research findings. Simmons, Nelson, and Simonsohn (2011), for example, find that even small changes in the choice of dependent variable in an analysis, or around the functional form of an analysis could lead to more than half of studies generating false-positive findings. Similarly, Strube (2006) finds that the practice of stopping data collection at a point where statistically meaningful findings have been found has the potential to increase the rate of false-positive findings.

The challenge, and debatably the trade-off, of facilitating exploration while maintaining academic, theoretical and statistical integrity, exists as a fundamental challenge for the statistical sciences (Gelman, 2019a). In the context of urban research, where the growth of data sets and research methods are seen as a source of opportunity for addressing climate and wider challenges in cities (Creutzig et al., 2019), these issues may be particularly large.

An illustration of the garden of forking paths is revealed in the dramatic differences between GHG consumption accounts for urban areas across literature. Across analyses that estimate consumption-based GHG emissions for large sets of cities (Moran et al., 2021; Nangini et al., 2019; Sun et al., 2022; Wiedmann et al., 2020) results regularly disagree by more than 50%. Key literature considering the factors leading to these divergent outcomes are identified in Owen’s work on the differences between multi-regional input output databases (2017; Owen, Steen-Olsen, Barrett, Wiedmann, & Lenzen, 2014, 2016) and Kilian, Owen, Newing, and Ivanova (2022)’s recent work on microdata selection.

6.2. Construct validity: are we testing what we think we’re testing?

Quantifiable observations enable the precision and objectivity of statistical approaches. The extent which such analyses provide insights, however, depends on the degree that hypotheses can be ‘operationalized’, or assessed by quantified data (Yarkoni, 2022). Where replication errors due to methodological or statistical mistakes are of a binary nature, Replication errors relating to the way hypotheses are operationalized are a matter of degree.

The construct validity of a measure is the extent that inferences can be made from the use of that measure as an operationalization of a hypothesis or concept. Many urban concepts to be engaged with by scholars are qualitative, disputed, and multidimensional, raising an intertwined set of challenges in this context. Capacity, for example, even as the concept is applied to climate change by researchers within the topic of urban transitions, seems to have as many definitions as there are publications (Castán Broto, Iwaszuk, & Westman, 2019; Filho et al., 2018; Gouldson et al., 2016; Wolfram, Borgström, & Farrelly, 2019). Social sustainability (Opp, 2017) and environmental sustainability (Purvis, Mao, & Robinson, 2019) are concepts widely addressed by researchers, but with a range of qualitative and quantitative measures. Nevertheless, a growing body of urban analysis boldly operationalizes urban concepts in order to apply
quantitative analysis. Grafakos, Trigg, Landauer, Chelleri, and Dhakal (2019), for example, develop an operationalization of the concept of ‘integrated climate action’ to assess climate action across a set of urban areas in Europe.

Two possible cases of generalizability error relating to the way hypotheses and concepts are operationalized can emerge. In the first case, a bold operationalization simply fails to appropriately capture a concept or idea. In the second, a concept or idea is captured, but the authors fail to define or incorrectly define the contexts in which their operationalization is relevant.

7. A framework for replication and generalizability

In Table 1, we define four cases to help distinguish between replication and generalizability. In this table, replication errors that emerge from outright statistical or methodological errors are not considered as they can occur in all research contexts.

The top-right presents the case where a phenomenon can be operationalized but varies across contexts making it hard to achieve content validity. In this case, a potential generalizability challenge looms larger than a challenge of replicability. A large number of technical and economic phenomena are characterized by this case: Income, energy use, GHG emissions and air pollution are all well-defined and quantified, and therefore operationalizable, but vary tremendously within and between urban settings, making achieving content validity more challenging due to the need for comprehensive datasets.

The top-left presents the case where the challenge of realizing content validity across contexts is relatively easier and the phenomenon can be highly operationalized. Technical, economic and social phenomena, where they have been shown to be relatively invariant between urban settings, have these characteristics. Dong, Santi, Liu, Zheng, and Ratti (2022) find scale-invariance of average commuting distances across urban areas in their study, building on wider literature that have confirmed common distributions of travel times across time, urban size and urban setting.

The bottom-right presents the case where a phenomenon is challenging to operationalize and the challenge of achieving content validity is high. Studies of phenomena with these characteristics, including of urban governance and civic pride, are presented with Table 1.

<table>
<thead>
<tr>
<th>Challenge of Replicability: Degree Phenomenon Can Be Operationalized</th>
<th>Easier to Operationalize</th>
<th>Example Phenomenon: Average Commuting Distance (Dong et al., 2022), the Amount of Energy to Meet Basic Human Needs (Millward-Hopkins, Steinberger, Rao, &amp; Oswald, 2020)</th>
<th>Example Phenomenon: Household Energy Use, Rebound Effects, Well-Specified Policy Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harder to Operationalize</td>
<td></td>
<td>Example Phenomenon: Effect of Public Engagement on the Quality of Urban Governance</td>
<td>Potential Challenge: Generalizability</td>
</tr>
<tr>
<td>Example Phenomenon: Civic Pride, Urban Governance, Capacity, Resilience, Co-Benefits of Climate Action</td>
<td>Potential Challenge: Generalizability and Replicability</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Replication errors that are the result of statistical mistakes of methodological errors are still possible in this research.
the greatest challenge across this framework as the phenomena vary across contexts and are challenging to operationalize. Studies of these phenomena are critical, however, to the advancement of the urban field as they are the foundation concepts of urban theory.

The bottom-left presents the case where it is relatively easier to achieve both content validity but operationalizing a phenomenon in challenging. Having certainty that content validity has been realized requires large numbers of comparable studies in order to assess whether a phenomenon and its context have been considered in sufficient detail. The comparability of studies, however, will be made challenging when operationalizations differ between authors.

Comparing studies that sit in different cells of Table 1 can help to think through the key challenges of advancing an integrated empirical urbanism. Well-developed studies of phenomena when datasets are available and comprehensive (which make content validity easier to achieve) can help to develop generalizable knowledge. Where datasets are more challenging to access and less comprehensive, however, small-scale studies will have more limited ability to speak across contexts. In these cases, well-developed case studies applying ideographic hypothesis are no less important towards understanding the generalizability of a phenomenon, as they can help to identify the aspects of the urban needed to capture subjects or phenomena in detail sufficient for more larger scale studies to achieve content validity.

Where phenomena are challenging to operationalize, small-scale studies play a different role. Disagreement over the appropriate means of operationalizing concepts such as governance, engagement, buy-in, and capacity, emphasizes the importance of large numbers of small-scale studies that can be compared to tease out the consequences of employing different operationalizations (and therefore understandings) of these concepts. Different approaches to operationalizing concepts also offer an opportunity to explore different understandings of the urban.

Table 1 also helps to illustrate the importance of identifying replication and generalizability as independent challenges. Where we confuse replicability with generalizability, we will confuse needing to consider the meaning of urban concepts such as governance or civic pride in specific contexts, with the need to consider a possible meaning of these and other concepts across urban areas. Conversely, when we confuse generalizability with replicability, we confuse needing to understand how concepts and metrics vary across urban spaces with the needed to understand urban concepts and metrics in detail in specific contexts.

8. Discussion

8.1. Boundary objects for a more coherent empirical urbanism

New empirical methods, some developed specifically for urban datasets and others developed in adjacent areas of research, are regularly developed to overcome the limitations of existing approaches. Agent-based models, for example, are heralded as providing a means of integrating disciplinary knowledge and overcoming disciplinary biases (Savin et al., 2023). Contextualized literature-based discovery with artificial intelligence is presented as radically decreasing the time taken to develop new approaches and ideas in the social sciences (Wang, Downey, Ji, & Hope, 2023).
The value these new methods offer should not be understated or reflexively dismissed. The new challenges introduced by these methods, however, may overshadow their contribution to urban knowledge. Contextualized literature-based discovery, for example, may suffer from understanding content but not context, creating novel and exciting combinations of urban ideas and concepts, but also combinations that are perplexing and misleading. The number of Centaurs may be indiscernible from the number of Frankensteins. More advanced agent-based modelling could amplify (rather than overcome) disciplinary biases when research teams are unable to bring together the right members. As a consequence of these and other challenges, new methods may sharpen inter and intra disciplinary urban divides. An application of contextualized literature-based discovery that develops a dozen fascinating combinations of urban concepts and one that is bizarre, may generate more attention for the single bizarre contribution than for the twelve fascinating contributions.

Critical to the advance of a more coherent empirical urbanism is therefore the development of methods and approaches that engage with both the epistemological as well as methodological challenges facing researchers. Applying replication and generalizability as boundary objects in combination with advances in methodological approaches may play a role in this context by providing a common, but flexible, point of reference for authors writing and reading empirical urban research.

Overarching the challenge of replicability is the question, ‘what are the characteristics of the urban’? Addressing this question requires thinking about the concepts and frameworks that seem appropriate in particular places and comparing them to those appropriate for other places. It also requires engaging with how urban areas are unique, the value and limitations of existing ways we differentiate between urban areas, and new ways of thinking about the ways urban spaces are fundamentally different.

Advances in analytical methods can support new ways of operationalizing urban concepts and testing urban theories when replication is applied as a conceptual focal point. For example, natural language processing can be applied to social media data to reveal sentiment and indicators of well-being. Clustering algorithms and data from transport applications can be used to consider spatial justice. Convolutional neural networks can be used to analyse patterns in urban green space usage to understand public health. In a study applying context-based literature discovery to urban literature, applying replication as a boundary object could help to emphasize that the success of such an analysis would not be in the number of new concepts created or their novelty, but in the extent to which these combinations challenge the way we currently characterize urban areas.

Overarching the challenge of generalizability is the question, ‘in what cases and contexts is place-based data and analysis indispensable? Alternatively, to what extent and under what conditions do ‘rules’ of cities exist, such that the specifics of an urban context are not important? A search for rules and axioms of the city jar with the fluidity, spontaneity, and messiness of the real-world urban that many researchers define as fundamental to the character of cities (Broto, 2020). At the same time, diversity, dynamic change and complexity do not necessarily preclude the existence of foundational relationships. Indeed, such relationships abound in the natural world (c.f Thommen et al., 2019)

To the challenge of generalizability, new empirical approaches offer means for capturing, stretching, and testing urban data when generalizability is applied as a conceptual
focal point. Machine learning approaches, including deep learning and neural networks, can be employed to find structures and patterns that researchers would have been unlikely to find on their own (Milojevic-Dupont & Creutzig, 2021). Machine learning enables capturing high-dimensional statistical patterns and complexities reflecting diversity in urban form, and, hence, can help to close the generalizability challenge. Resampling and simulation-based approaches, including bootstrapping and permutation testing, can help to test dependencies between predictors and correct for biases. Boosting, stacking and Bayesian model averaging allow researchers to combine inferences from multiple datasets (Gelman & Vehtari, 2021). These methods, among others, offer the opportunity to take the conceptual starting points offered by existing literature (e.g. Creutzig, 2015; Eisenack & Roggero, 2022; Hsu et al., 2020) and expand and deepen the analysis. In a study applying advances in agent-based modelling to urban analysis, for example, applying generalizability as a boundary object could help to emphasize that the importance of such analysis lies less in adding to the set of factors that are common across cities and more in the way these findings support or refute theorizations.

8.2. Boundary objects and the development of a coherent interdisciplinary urbanism

Boundary objects, including replication and generalizability, may be valuable for advancing a coherent empirical urbanism. The role of boundary objects in advancing the urban field more generally, however, will be more complicated owing to the layered complexity of the urban as a subject.

Cities loom large as sources of environmental harm. But in the wealthy world the city as a forest of smokestacks sending dirt and smog over surrounding nature has been replaced by the city as a node of consumption, drawing resources from its hinterland and causing destruction far from its borders. In the wealthiest cities emissions associated with consumption of goods and services can be more than two times the emissions associated with energy use (Sudmant, Gouldson, Millward-Hopkins, Scott, & Barrett, 2018).

The city is increasingly seen as a driver of political, social and technical change. Economists have long proselytized about cities and innovation (Fujita, Krugman, & Venables, 1999). Indeed, the process of urbanization itself is a form of technology for addressing some of the sustainability challenges we face: bringing people closer creates opportunities for the sharing infrastructure and services (Ottelin, Heinonen, Nässén, & Junnila, 2019). Cities are also sites for experimentation in governance (Dignum, Dorst, van Schie, Dassen, & Raven, 2020; Hölscher & Frantzeskaki, 2020) and leadership in the setting of climate action commitments (Howarth, Lane, & Fankhauser, 2021). At the same time, the city is also the ultimate representation of irreversible change, the place where social, technical and political lock-ins are manifest, and the physical demonstration of humanity’s short-sighted domination of nature in the Anthropocene (Brondizio et al., 2016; Bulkeley, 2021; Bulkeley, Drew, Hobbs, & Head, 2018; Keys et al., 2019).

In these ways and others, the city is thus the physical and metaphorical place where the messy contradictions that characterize the twenty-first century become most apparent. Contemporary urban research and the project of an urban science have to speak to both explosive growth – in terms of population, infrastructure, area covered – and other metrics. Stagnation and decline. Staggering inequality and broad-based human
flourishing. Appalling waste and astounding efficiency. Dynamism, change, lock-in and rigidity. Rising diversity at a local scale and increasing homogeneity at a global scale. Dependencies between places that support resilience and that simultaneously increase risks.

Where frameworks of North/South, urban/rural, centre/periphery, and primary/secondary/tertiary once offered insight, the patchwork urban of the twenty-first century is not only unique, but unique in ways we do not yet understand (Brenner & Schmid, 2015). Even as the term ‘urban’ has become common as a descriptor for the world we live in and the challenges we face, what can be defined as ‘urban’ has a more fluid and harder to pin down meaning (Wachsmuth, 2014).

These complexities of the urban subject are both the case for interdisciplinarity and the cause of a siloing of disciplinary approaches. Applying specialized knowledge to engage with urban questions not only abets further specialist knowledge but can deepen disciplinary-specific norms, values, identities and languages. Depth of urban knowledge can be at the cost of breadth. And since many of the disciplines considering urban topics are physically and metaphorically housed in the departments of other fields of research, including economics, biology, geography and increasingly computer science (Ball, 2002; Barnes, 2013; Purvis, 2021; Sennett, 1996), bridging not just methods and practices but ‘social entities’ will require a different approach to the use of boundary objects.

Where applied between disciplines with a substantial set of shared characteristics, boundary objects can be designed to leverage these shared characteristics. Empirical and computational biologists use many of the same methodological approaches as computer scientists and physicists. For authors from these fields, replication, generalizability and other boundary objects can be applied as conceptual meeting points for literatures that have applied similar methods but may use slightly different languages and may not have been aware of literature in other disciplines considering the same questions.

When applied between disciplines with few shared characteristics, boundary objects need to be more abstract, functioning as conceptual scaffolds to support the translation of methodologies, findings, and theories between disciplines. Used this way, boundary objects can enable a dialogue where the intricacies of one field can inform and enrich the research questions and hypotheses of another, allowing for collaborative exploration that extends beyond the conventional boundaries of each discipline.

9. Conclusion

The global challenges cities play a role in are too urgent for opportunities for action to be learned and relearned on a city-by-city basis. But how can analyses speak to cities over any meaningful scale when the information we have is nearly always incomplete and key urban concepts defy being represented in straightforward ways? Moreover, how can knowledge advance when subject matter covers multiple disciplines? Embedding replication and generalizability as boundary objects in the way empirical research questions are posed and results are presented can contribute to the advancement of urban knowledge in ways that are prosaic, and in some cases, (possibly) profound.

The concept of replication serves as a meeting point for research that considers the nature of urban phenomena. This makes the concept of replication more readily
applicable to qualitative research, ideographic hypotheses, and subjective and highly contextual research. Phenomena that are effectively operationalized, however, can then be tested for the extent they generalize across urban contexts. Embedding the concept of replication is therefore valuable even when researchers are most interested in generalizability.

The concept of generalizability serves as a meeting point for research that is considering when urban phenomena transcend the particularities of individual cases. This makes the concept of generalizability more readily applicable to empirical research and nomothetic hypotheses. Individual cities and cases that break sharply with the empirical relationships that hold across other cases and cities, however, can suggest gaps in our understanding of urban concepts and highlight a need for replication research. Indeed, the extent to which all observations in empirical analyses invariably fail to conform strictly to defined patterns and derived relationships suggests possible opportunities for deeper learning. Embedding the concept of generalizability in empirical research is therefore also critical for research focusing on replication.

More profoundly, replication, generalizability and boundary objects more generally, may also contribute to enabling interdisciplinary urbanism. Many disciplines share common methods and concepts while developing research for distinct sets of academic journals and rarely engaging with one another’s work. Boundary objects in these cases are reference points that can help to leverage breakthroughs and findings between fields. Where disciplines are more distinct, boundary objects act as ‘trading zones’, providing shared objects to study around which researchers can align their distinct analytical tools and social theories (Galison, 1999). Testing and critique of replication and generalizability applied in this role may be important for advancing the urban field.

Notes

1. Cities are not synonymous with urban areas. For the purposes of this manuscript; however, the two terms are used interchangeably.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

A. Sudmant http://orcid.org/0000-0001-8650-8419

References


Gelman, A. (2019b). When we make recommendations for scientific practice, we are (at best) acting as social scientists. European journal of clinical investigation, 49(10), e13165.


Schuurman, N., & Pratt, G. (2002). Care of the Subject: Feminism and Critiques of GIS. *Gender, Place & Culture*, 9(3), 291–299. [https://doi.org/10.1080/096636902200003905](https://doi.org/10.1080/096636902200003905)


Wachsmuth, D. (2014). City as Ideology: Reconciling the Explosion of the City Form with the Tenacity of the City Concept. *Environment and Planning D: Society and Space*, 32(1), 75–90. [https://doi.org/10.1068/d21911](https://doi.org/10.1068/d21911)


