

Article

Adaptive Governance Through Urban Regulatory Sandboxes: Expert Perceptions from Four Spanish Climate-Neutral Cities

Ana Horta-Bellido ^{1,*}, Manuel Alméstar ², Elisa Navarro-Carrillo ^{3,*}, Cristina Armuña ¹ and Claudio Feijoo ¹

¹ Escuela Técnica Superior de Ingenieros Industriales de Madrid, Universidad Politécnica de Madrid, 28006 Madrid, Spain; cristina.armuna@ruleeleven.es (C.A.); claudio.feijoo@upm.es (C.F.)

² Innovation and Technology for Development Centre—itdUPM, Universidad Politécnica de Madrid, 28040 Madrid, Spain; manuel.almestar@upm.es

³ Centro de Apoyo a la Innovación Tecnológica (CAIT), Universidad Politécnica de Madrid, 28223 Pozuelo de Alarcón, Spain

* Correspondence: ana.horta@alumnos.upm.es (A.H.-B.); elisa.navarro@upm.es (E.N.-C.)

Abstract

Cities face increasing pressure to accelerate climate action while operating within governance systems often characterised by institutional rigidity and slow regulatory adaptation. In this context, urban regulatory sandboxes have emerged as experimental instruments that enable municipalities to test innovative solutions while adapting institutional practices under conditions of uncertainty. Despite their growing implementation, limited evidence exists regarding their capacity to transform municipal governance itself. This research examines the conditions under which urban regulatory sandboxes can foster the governance capacities required for municipalities to lead the climate transition. A sequential mixed-methods design was applied, combining semi-structured interviews with municipal practitioners and a Likert-scale self-assessment instrument completed by the same practitioners. Four Spanish cities with active urban sandboxes—Barcelona, Madrid, Valencia, and Valladolid—were analysed across three governance dimensions: learning vs. isolation, flexibility vs. rigidity, and climate mission alignment. The findings suggest that, in the perception of the practitioners interviewed, urban sandboxes foster adaptive governance logics, though institutionalising these capacities exposes real tensions around intra- and inter-organisational collaboration, and the depth of transformation differs by each city context. These results carry practical implications for municipal policymakers seeking to integrate sandboxes as tools for climate-responsive urban governance and as enablers of local entrepreneurship and innovation ecosystems. This study contributes to the literature on adaptive governance and urban experimentation by positioning urban regulatory sandboxes as emerging instruments for leading climate transitions under conditions of uncertainty. Ultimately, urban sandboxes can serve as catalysts for embedding climate adaptability within municipal governance but only when accompanied by the deliberate institutional work needed to sustain them.



Academic Editors: Jianming Cai and Giuseppe Ioppolo

Received: 24 July 2026

Revised: 12 September 2026

Accepted: 17 September 2026

Published: 24 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: regulatory experimentation; multilevel governance; policy innovation; innovation ecosystems; public entrepreneurship

1. Introduction

Cities have become the principal arena for implementing climate action. While national and supranational institutions define policy frameworks, targets, and regulatory conditions, municipalities are responsible for translating these commitments into concrete

interventions. It is at this scale that ambitious climate objectives encounter the administrative, institutional, and governance challenges of implementation [1,2]. Municipal governments are expected to deliver on binding decarbonisation commitments through administrative systems traditionally designed to prioritise stability, procedural predictability, and sectoral specialisation rather than rapid institutional adaptation [3,4]. This tension between the urgency of the climate transition and the institutional mechanisms of municipal administration has fuelled a growing interest, in both research and policy practice, in instruments that allow cities to govern differently: to test, learn, and revise policy under conditions of genuine uncertainty rather than applying fixed rules to problems that are still being defined [5].

Against this backdrop, urban regulatory sandboxes have emerged as an experimental instrument intended to create protected spaces for regulatory experimentation, institutional learning, and policy adaptation under conditions of uncertainty. Originally developed in the financial sector as bounded spaces in which firms could trial novel products under relaxed supervisory conditions [6], the sandbox model has progressively expanded into urban governance, where it is increasingly used to test climate-related technologies, services, and business models while temporarily adapting or exempting specific regulatory requirements under controlled conditions [7].

In an urban context, regulatory sandboxes are intended to enable municipalities to experiment with climate-related technologies, services, and business models while temporarily adjusting specific regulatory procedures in order to facilitate innovation and generate institutional learning. Within initiatives such as the EU Mission for Climate-Neutral and Smart Cities (Cities Mission), they are increasingly promoted as mechanisms capable of transforming regulation itself into an object of experimentation, allowing cities to learn not only about the solutions being tested but also about the governance arrangements required to deploy them at scale [8]. These ambitions ultimately rest on the assumption that regulatory experimentation can strengthen the governance capacities required for municipalities to respond more effectively to the climate transition. However, empirical evidence remains limited on whether urban regulatory sandboxes actually foster these governance capacities or whether they primarily function as symbolic or isolated experimentation mechanisms.

This question extends beyond the instrumental performance of urban regulatory sandboxes. It points to a long-standing debate in the governance literature on the difference between formal and substantive institutional flexibility, and whether experimentation conducted at the local level can lastingly reconfigure the governance-supporting architecture or instead remains confined to isolated, non-replicated pilots [9]. Part of the empirical literature on urban climate experimentation has documented precisely this risk: individual experiments that succeed on their own terms while leaving the broader institutional environment essentially unchanged, disconnected from formal planning and mainstream urban governance processes. This is a pattern described in the literature as a “pilot paradox” [4,10,11]. Whether urban regulatory sandboxes escape this paradox, and under what conditions, is the empirical problem this study addresses.

The Spanish case provides a valuable setting in which to examine this question. Spain’s participation in the European Mission for 100 Climate-Neutral and Smart Cities has led, within only a few years, to the emergence of several municipalities that have negotiated Climate City Contracts (CCC) and, under the national legal framework established by Royal Decree 568/2022, have begun to implement their own urban regulatory sandboxes [12]. This provides an opportunity to observe the same governance instrument across cities with markedly different administrative traditions, institutional capacities, and political contexts while operating under a common national and European framework.

Against this background, this study addresses the following research question: to what extent, and under what conditions, do urban regulatory sandboxes foster the governance capacities that municipalities need to lead the climate transition? Specifically, it analyses four Spanish cities with active urban sandboxes—Barcelona, Madrid, Valencia, and Valladolid—using three complementary dimensions of adaptive governance: learning versus isolation [13–17], flexibility versus rigidity [18–22], and climate mission alignment [23–26]. A sequential mixed-methods design combining semi-structured interviews with municipal practitioners and a Likert-scale assessment instrument is employed to examine how the same governance instrument performs across different municipal contexts. Drawing on the adaptive governance literature, these three dimensions are conceptualised as complementary governance capacities that enable municipalities to learn, coordinate, and adapt institutional responses to the climate transition. In doing so, the study seeks to understand not only whether urban regulatory sandboxes foster these governance capacities, but also how their nature and extent vary once a common policy instrument is embedded within distinct local governance architectures.

The structure of the article is organised as follows. Section 2 develops the theoretical framework, placing urban regulatory sandboxes at the intersection of adaptive governance, mission-oriented innovation policy, and multi-level governance theory, and defining the analytical dimensions used in the empirical analysis. Section 3 presents the materials and methods, including the case selection strategy and the mixed-methods design. Section 4 reports the results for the governance dimension. Section 5 discusses these findings in light of the theoretical framework, and Section 6 concludes with implications for research, limitations faced and municipal practice.

2. Governance Capacities for the Climate Transition: Adaptive Governance and Urban Regulatory Sandboxes

Urban regulatory sandboxes offer cities a strategic instrument for testing climate-neutral policies under conditions of relevant uncertainty, allowing municipal governments to relax, reinterpret, or temporarily suspend standard regulatory procedures in pursuit of experimentation and learning [27]. Municipal governance has traditionally been structured around rigid administrative procedures, sectoral fragmentation between departments, and limited mechanisms for institutional learning across levels of government [28]. However, the urgency of the climate transition, together with supranational frameworks such as the EU Mission for 100 Climate-Neutral and Smart Cities, is pressuring public administrations to adopt more adaptive, flexible, and multi-level approaches capable of responding to fast-changing environmental and technological conditions [29]. Whether this adaptive capacity translates into effective institutional flexibility, however, depends on how sandboxes are conceived, by whom, and within what governance architecture.

The subsections that follow build this argument in three steps: adaptive governance depends on institutional capacities being present and sustained; the sandbox is the mechanism through which cities try to build those capacities by reconciling mission-driven alignment with adaptive learning; and this rarely works through the sandbox alone, since flexibility ultimately depends on the wider multi-level architecture it sits within.

2.1. Adaptive Governance and Urban Climate Experimentation

Adaptive governance emerges in the literature on social-ecological systems: institutions capable of learning and adjusting as conditions change, rather than relying on rules fixed in advance [30,31]. Boyd and Juhola [32] extend this to the urban context, treating climate governance as a continuous process of trial, observation, and revision rather than a one-off regulatory act. This shift is consistent with broader bibliometric evidence showing

the urban resilience field moving from a narrow, engineering-based focus on disaster prevention toward multidimensional, socio-ecological-economic adaptive systems increasingly organised around multi-level governance [33].

A parallel strand of the literature emphasises institutional innovation and learning as preconditions for moving beyond reactive policy, pointing to a shift from fragmented environmental management toward integrated, cross-departmental experimentation [34]. This rarely requires formal reform: evidence from Madrid shows that institutional flexibility achieved through the reinterpretation of existing rules redistributes decision-making power without changing the underlying legal framework [35]. A comparable dynamic appears in urban environmental governance elsewhere: in China, the transition of the Waliu Community's vending zones toward spatial self-management redistributed governance authority to residents without any formal regulatory change [36]. Whether such innovations get absorbed into institutional practice or remain isolated is an empirical question that reflects the tension this study finds between knowledge-sharing and resistance across the four municipalities.

This tension between local success and systemic absorption is what van Buuren et al. [11] term the “pilot paradox”: drawing on Dutch climate adaptation projects, they show that the same conditions that let a pilot succeed on its own terms—protection from standard procedures, dedicated resources, motivated participants—are often precisely what keep it disconnected from the mainstream governance processes it was meant to inform. Internal success factors and external, systemic success factors can therefore pull in different directions, and a pilot can score highly on the first while contributing little to the second [11]. Evans, Karvonen and Raven [10] document the same disconnection at scale across urban experimentation more broadly. This distinction, between a sandbox's internal performance and its external, systemic effect on municipal governance, structures the analysis that follows, and is returned to directly in Section 5.2. Yet large-scale urban experimentation databases suggest this goal is only partially achieved. Experiments provide little explicit ecological knowledge, are led primarily by local authorities, and show limited evidence of planned adaptation [32]. This gap between ambition and practice points to the institutional capacities that determine whether adaptive principles are actually operationalised.

Gupta et al. [37] capture this through the concept of governance capacities: learning capacity, room for autonomous change, leadership, and resource availability, dimensions that determine whether an institution can actually adapt rather than merely claim to. Adaptive governance is thus a matter of degree, not a property a city simply has or lacks. It is, then, an ongoing achievement rather than a status attained once and for all, dependent on specific institutional capacities being present and sustained over time. This is what motivates closer attention to the mechanisms, such as urban sandboxes, through which cities attempt to build and exercise those capacities in practice.

2.2. Regulatory Sandboxes and Mission-Oriented Urban Governance

A regulatory sandbox is a bounded space, geographic or sectoral, in which innovators test new products or business models under a tailored supervisory framework, usually in exchange for temporary exemptions from specific rules [38]. First formalised by the UK Financial Conduct Authority in 2015, the instrument has since spread beyond fintech into artificial intelligence, energy, and industrial policy [39,40]. This is an instance of what Park [41] calls regulation for innovativeness: rules designed to enable experimentation without abandoning administrative focus. The same exposure to real institutional conditions also draws entrepreneurs to the sandbox, giving startups a structured route into the city, and municipalities a structured way to engage with them rather than case by case [42]. This expansion has accelerated further since 2024, as regulatory sandboxes have been

adapted to increasingly complex domains. In the energy and mobility sector, Wady and Consoni [43] show how a sandbox for electric-vehicle charging infrastructure can generate the tariff and regulatory certainty needed for infrastructure investment, illustrating how far the instrument now travels beyond its financial origins. Kálmán's [44] comparative case study of fintech sandboxes across the UK, Singapore, and Hungary makes a related point from a different angle: mature and emerging sandbox ecosystems differ less in their formal design than in the surrounding institutional capacity to absorb what a sandbox produces—a comparative lesson that speaks directly to the asymmetries this study documents across Spanish municipalities (Section 5.4). Most directly relevant to the present study, however, is Okonjo's [45] analysis of AI regulatory sandboxes adapted for public sector experimentation: because conventional sandboxes were designed for private market entry rather than public sector institutionalisation, adapting them requires, in Okonjo's words, "substantial normative, legal and institutional adaptation"—precisely the kind of adaptation this study finds only partially achieved when sandbox-generated knowledge is not yet embedded in standing municipal practice.

This approach meets mission-oriented innovation policy in the urban domain, where public resources are mobilised around bold, time-bound goals [46,47]. The EU's Mission for 100 Climate-Neutral and Smart Cities illustrates this by using the formula of CCC. This binds municipalities, industry, academia, and civil society to shared commitments [7]. These frameworks may hold together at the national level, but they tend to cause friction once cities, with their many actors and limited resources, have to put them into practice [48].

The sandbox sits at that friction point. If mission-oriented governance supplies strategic alignment and adaptive governance supplies iterative learning, the urban sandbox is the procedural mechanism through which cities try to reconcile both and, by doing so, build the governance capacities that determine whether adaptive governance holds in practice. Whether that promise is met, or the sandbox reproduces the rigidities it was meant to overcome, is addressed in the next section.

2.3. Institutional Flexibility, Rigidity, and Multi-Level Governance

Sandboxes rarely deliver flexibility by themselves. Whether an instrument actually reshapes practice, or simply sits alongside it, depends on the institutional structure surrounding it. Hooghe and Marks [25] distinguish Type I governance—general-purpose, durable, non-overlapping—from Type II, task-specific and flexible. Municipal climate governance, including the urban sandbox, sits awkwardly between the two: nested in the durable architecture of municipal administration, yet meant to function as a bounded, task-specific instrument. Much of the friction documented in the sandbox literature can be traced back to this hybrid position.

Moyo et al. [29] push this further by identifying four recurring misfits between national strategy and local implementation: institutional, strategic, financial, and epistemic. Their point is that a system can look formally adaptive and decentralised while remaining, in substance, poorly matched to local conditions. That gap between formal and substantive flexibility surfaces repeatedly in the literature.

Park [41] offers a case where the gap narrows: sandbox-driven innovation in South Korea owes less to the instrument itself than to sustained government support and institutional learning, what Park calls "Dynamic Regulatory Learning". Flexibility, on this evidence, comes from institutional capacity rather than from the sandbox mechanism. This is corroborated by Kim and Cho [49], who find that in Korea's regulatory sandbox scheme, government policy orientation (progressive vs. conservative) systematically moderates how quickly urban sandbox-generated learning is absorbed into regulatory reform. This is

evidence that institutional and political context, not the sandbox format itself, determines whether adaptive capacity is realised.

3. Materials and Methods

This study adopts a qualitative comparative research design to examine the extent to which the implementation of urban regulatory sandboxes transforms traditional governance models into more adaptive structures capable of addressing the climate emergency. The analysis is structured around three interrelated analytical dimensions: learning vs. isolation (GOV_D1), flexibility vs. rigidity (GOV_D2), and climate mission alignment (GOV_D3). These dimensions are conceptualised as complementary aspects of institutional adaptation: GOV_D1 captures whether sandbox-generated knowledge circulates and consolidates across departments and territories or instead remains confined to isolated pilots; GOV_D2 evaluates whether municipal governance achieves substantive flexibility in practice or merely displays its formal appearance; and GOV_D3 assesses the degree to which urban sandbox operation is explicitly aligned with supranational and national climate missions rather than pursued as an end in itself. These three dimensions are conceptualised, and empirically assessed throughout this study, as governance capacities perceived and enacted by the officials directly responsible for each sandbox, rather than as capacities independently verified against objective institutional records—such as formal governance reforms, budgetary allocations, meeting minutes, or external evaluation reports. Such records were not available in a form comparable across the four municipalities at the time of data collection, and their absence is revisited as a scope limitation in Section 6.3.

3.1. Materials: Spanish Multi-Level Governance and the Climate-Neutral Cities Mission

Spain's territorial governance follows a three-tier structure in which the central state, the autonomous communities, and municipalities hold overlapping but constitutionally distinct competences, producing precisely the kind of nested, general-purpose architecture that Hooghe and Marks [25] describe as Type I governance. Climate and urban planning competences are distributed asymmetrically across the system, which means that any municipal-level instrument, including an urban sandbox, must be read against a background of shared and sometimes contested authority rather than as an autonomous local decision.

Within this structure, the European Commission's Mission for 100 Climate-Neutral and Smart Cities provided the entry point through the negotiating of a CCC, formalising binding climate commitments in coordination with national and regional authorities [7]. Spain's participation in the Mission was first defined by the selection of seven cities. Of these, a subset formalised CCC as described above, and of that subset, Madrid, Barcelona, Valencia, and Valladolid are the four that have an approved and operative urban regulatory sandbox. This funnel, national Mission cities, the CCC and then cities with a functional sandbox, defines the sample analysed in this study.

At the national level, Royal Decree 568/2022 established a general legal framework for regulatory sandboxes [12], creating the statutory basis on which municipalities could, in principle, design locally adapted experimentation mechanisms without requiring national legislation for each case. Whether this national framework is interpreted and implemented consistently across municipalities with significantly different administrative capacities, political priorities, and prior experience with collaborative planning, as documented for Madrid [35], is an open empirical question rather than something the legal framework itself resolves.

It is this combination—a supranational mission, a national regulatory anchor, and four municipalities operating under broadly similar formal constraints—that frames the comparative analysis developed in the remainder of this study.

3.2. Methods

The case study was adopted as the main methodological approach of this research, given its value for analysing complex social phenomena where cause-and-effect relationships are not readily observable and are strongly shaped by context [50,51]. The case study is considered particularly suitable for theory-building in emerging fields, as it identifies and articulates relevant variables and mechanisms based on empirical observation [52]. Accordingly, this study is explicitly positioned as exploratory and theory-building rather than confirmatory: given how recently urban regulatory sandboxes have been adopted by Spanish municipalities, the population of practitioners with direct, first-hand responsibility for their operation is itself small, and the sample analysed here should be read as a purposeful pilot sample of that population rather than as an undersized draw from a much larger universe of comparable cases. A mapping of cities that have adopted or are in the process of adopting urban sandbox frameworks was conducted to define the scope of the research. Between 2022 and 2026, at least eleven Spanish cities have developed or initiated regulatory frameworks for urban sandboxes, reflecting a rapid process of policy diffusion across different territorial and institutional contexts.

A two-stage filtering strategy was applied to define the final sample. The first filter relates to strategic alignment with European climate policy, specifically the EU Mission for Climate-Neutral and Smart Cities: only municipalities that are either officially part of the Mission or have obtained the “Mission City” label were considered. The second filter is methodological and relates to the operational status of the urban sandbox: given the study’s mixed-methods approach, only cities where the sandbox is currently active were included. Applying both filters resulted in a final sample of four case studies: Madrid, Barcelona, Valencia, and Valladolid (see Figure 1). These cities represent different stages of institutional development and urban scale while sharing a common orientation towards climate neutrality and innovation-driven policy experimentation.

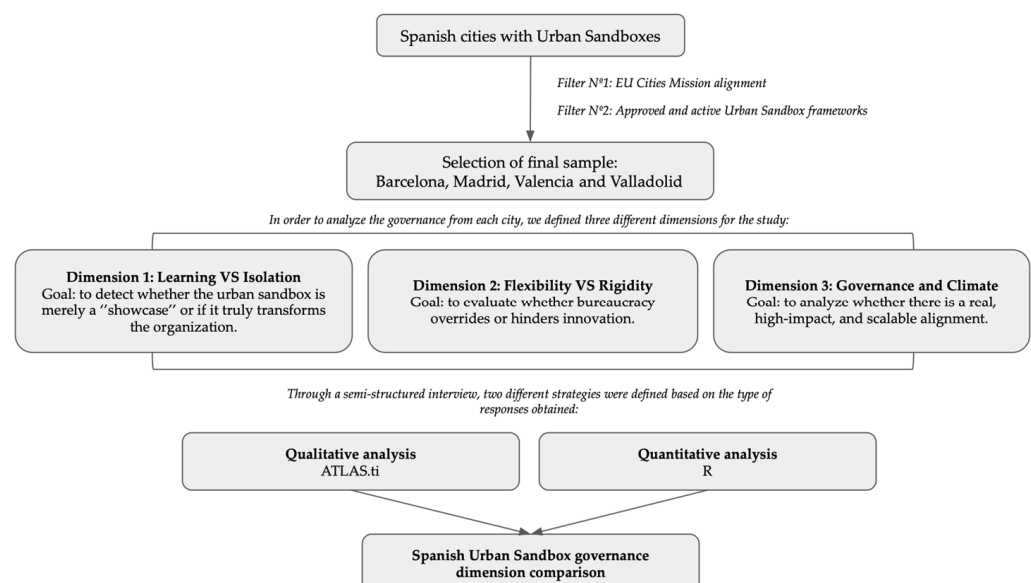


Figure 1. Sample selection and mixed-methods analytical process across the four case studies. Arrows indicate the sequential flow of the research design, converging in the final comparative analysis. (Source: own elaboration).

Six semi-structured interviews were conducted with sandbox representatives across the four municipalities (three in Valencia, and one each in Barcelona, Madrid, and Valladolid), reflecting differences in institutional structure and the number of actors directly responsible for sandbox operation in each city. In every city, the official with overall responsibility for the urban sandbox was interviewed; in Valencia, the most institutionally developed of the four sandboxes, two additional technical staff directly carrying out its day-to-day management were also interviewed. This addition follows a key-informant sampling logic proportional to each case's institutional maturity: the larger operational team behind Valencia's urban sandbox justified capturing more than a single perspective, whereas in the other three cities the lead official was also the primary source of first-hand operational knowledge. It is also worth noting that, as the officials with direct responsibility for each sandbox's operation, all respondents have a vested interest in its perceived success—a further source of potential bias that complements the social-desirability concern addressed in Section 3.2.1 and is revisited as a limitation in Section 6.3. Interviews were conducted online and lasted between 50 and 90 min.

Table 1 summarises the institutional design of each sandbox in terms of the legal instrument adopted, application frequency, target scope and participants, and evaluation committee structure, providing the institutional backdrop against which the governance dimensions analysed in this study are interpreted.

Table 1. Comparative instrumental design of the Urban Sandboxes by City Council. (Source: own elaboration).

City Council	Legal Instrument	Application Frequency	Target Scope and Participants	Evaluation Committee
Barcelona	Municipal ordinance.	Year-round, open.	Any municipal-competence theme; housing emphasis; startups, universities and research centers.	Standing members and ad hoc experts by topic.
Madrid	Municipal ordinance.	By calls.	Sectors per call (3 of 9 verticals); TRL between 3 and 8, non-commercial stage, open to all profiles.	Innovation Councillor (chair) and area directors-general; ad hoc experts by topic.
Valencia	Municipal ordinance.	Year-round, open.	Sector-agnostic; TRL around 5, non-commercial stage, open to all profiles.	Fixed core and host-service experts; external experts inclusion are optional.
Valladolid	Municipal ordinance.	Year-round, open.	Sector-agnostic; Startups firms, tech centers, universities and consortiums.	Two tiers: IdeVa staff with partners and then for execution, municipal areas.

3.2.1. Data Collection

Interviewees held positions directly related to municipal innovation management: two were innovation officers (33.3%), one was a head of service (16.7%), and the remaining three (50%) comprised an innovation manager, an innovation coordinator, and a technical advisor, one each. Most respondents were attached to government areas linked to innovation or, alternatively, directly to the municipal management office or the mayor's office, reflecting a deliberate sampling focus on actors with operational and managerial responsibility over the sandbox in their respective cities.

The interview protocol was designed as a semi-structured instrument combining Likert-scale, closed, and open-ended questions, included in Appendix A. A six-point Likert scale (1 = strongly disagree; 6 = strongly agree) was used to avoid neutral responses and

encourage respondents to position their assessments clearly. Because these items were self-assessed by the same officials responsible for managing the urban sandbox, the research is potentially exposed to social desirability bias, i.e., a tendency to rate one's own performance more favourably than an independent observer would. To mitigate this, each Likert item was paired with open-ended follow-up questions requesting concrete examples and explanations, and quantitative scores were subsequently triangulated against the qualitative narratives obtained (see Section 3.2.2); convergence between the two strands was treated as strengthening a finding, while divergence was treated as a signal of possible inflation rather than as noise. This strategy is applied in Sections 4.3 and 5.3, and the residual risk is revisited as a limitation in Section 6. Closed categorical questions were additionally included to capture specific governance arrangements that do not lend themselves to a Likert format: internal resistance to integration, the existence of a national intermunicipal exchange mechanism, and the existence of an international exchange mechanism.

The protocol covered seven Likert items distributed across the three governance dimensions described above (GOV_D1, GOV_D2, GOV_D3). For Valencia, where three respondents were interviewed, item-level scores were averaged across the three; for the closed categorical questions, the response of the most senior official was used as the reference value.

3.2.2. Data Analysis

Following a sequential mixed-methods design, qualitative and quantitative data were analysed separately and then integrated at the interpretation stage as indicated in Figure 2. Interview transcripts were analysed in ATLAS.ti (ATLAS.ti Scientific Software Development GmbH, Berlin, Germany; version 26.0.1) following a thematic coding approach, with the codebook organised around the three governance dimensions and refined inductively through iterative readings of the transcripts. Relational networks were subsequently constructed in ATLAS.ti to visualise connections between codes and identify patterns of co-occurrence across the cities interviewed. Within this semantic network, knowledge sharing emerged as the most connected node, while the only contradictory relation identified linked interdepartmental collaboration to internal resistance to integration.

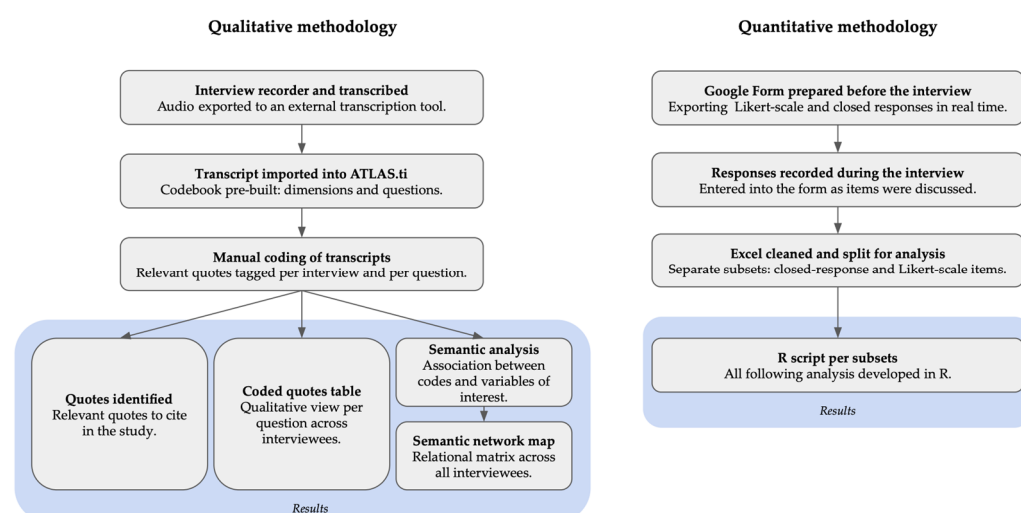


Figure 2. Qualitative and quantitative data analysis process. Arrows indicate the sequential flow of the analysis process, including the branching of manual coding into its three qualitative outputs. (Source: own elaboration).

Quantitative responses obtained through the Likert-scale assessment were analysed using descriptive statistics in R (The R Foundation for Statistical Computing, Vienna, Aus-

tria; version 4.5.2). For each item and dimension, means and standard deviations were calculated both globally and disaggregated by city. Given the small sample size ($n = 6$, with $n = 1$ per city except Valencia, $n = 3$), no inferential statistical testing was applied; the analysis instead relies on descriptive statistics and on between-city variation, and all interpretive claims are framed in tentative, descriptive terms throughout the remainder of this study. This analytical choice is consistent with the exploratory, theory-building purpose of the study outlined in Section 3.2: rather than testing pre-specified hypotheses on a representative sample, the aim is to generate empirically grounded propositions about adaptive governance in urban sandboxes that future research can test more formally by drawing on a larger population as the phenomenon matures. Integration of the quantitative and qualitative strands was achieved through a convergent triangulation strategy, systematically comparing descriptive results with qualitative narratives for each city and governance dimension; convergent findings were treated as more robust, while discrepancies were examined as substantively meaningful rather than as methodological inconsistencies. Generative AI (Claude, Anthropic; Sonnet 4.6) was used to assist with the development and debugging of the R scripts employed for this quantitative analysis. All code outputs were reviewed, tested, and validated by the authors, who take full responsibility for the analysis and its results.

4. Results

This section presents the results obtained from the qualitative and quantitative analysis of the information gathered through the semi-structured interviews, focusing on the Governance dimension of the active urban sandboxes operated by Spanish cities linked to the EU Cities Mission. The limited number of interviews conducted ($n = 6$), three from Valencia City Hall and one each from Barcelona, Valladolid and Madrid City Halls, means that the findings should be interpreted as descriptive rather than statistically generalisable.

4.1. Qualitative Results: Thematic Analysis of Stakeholder Interviews with ATLAS.ti

As shown in Table 2, the governance code group accounts for 57 quotations across the six interviews, distributed unevenly among the responses: Madrid City Hall contributed the largest share (13 quotations), followed by Valladolid (12), while Barcelona and one of the additional Valencia interviews each contributed the smallest share (6). Taken together, the three from Valencia account for 26 quotations, a volume comparable to Madrid and Valladolid combined.

Table 2. Distribution of quotations by code group and document for the Governance dimension.

Barcelona City Hall	Madrid City Hall	Valencia City Hall (1)	Valencia City Hall (2)	Valencia City Hall (3)	Valladolid City Hall	Total
6	13	11	9	6	12	57

Figure 3 presents the semantic network generated in ATLAS.ti, where nodes represent codes and links the relations established between them (is the cause of, is associated with, and contradicts). The network is organised around two connected clusters. The first follows a causal sequence in which regulatory flexibility (Dimension 2) is linked as a cause of systematisation, which in turn is a cause of knowledge sharing (Dimension 1); regulatory flexibility is also linked as a cause of interdepartmental collaboration and of risk tolerance, and is itself associated with response time. The second cluster centres on knowledge sharing, reached through a causal chain from multi-level coordination (Dimension 3), and associated with both intermunicipal exchange and international exchange (Dimension 3). A single contradictory relation links interdepartmental collaboration and internal resistance

(Dimension 1), while response time remains the most peripheral node, connected only to regulatory flexibility.

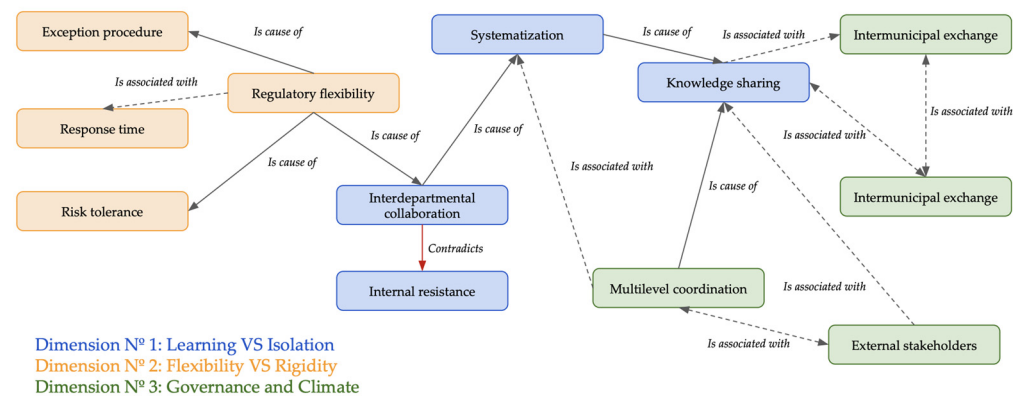


Figure 3. Semantic network of the Governance code group. Solid arrows indicate causal relations (“is cause of”), dashed arrows indicate associative relations (“is associated with”), and the red arrow indicates a contradictory relation (“contradicts”); node colours correspond to the three governance dimensions. (Source: own elaboration, ATLAS.ti).

4.2. Qualitative Results: Categorical Evidence

Turning to the categorical evidence, the responses to the three closed questions are summarised by city in Table 3, where each cell is coloured according to the response category. Agreement across the four city councils was absent across all three items. Regarding internal resistance to integrating sandbox results into municipal operations, Valencia and Valladolid answered “Yes”, Barcelona answered “No”, and Madrid reported the situation as “Unknown”. A comparable divergence appears in the exchange mechanism with other municipalities, confirmed (“Yes”) by Valencia, Barcelona and Valladolid but reported as “In progress” by Madrid. The international exchange mechanism constitutes the most polarised case: Valencia and Valladolid again answered “Yes”, while Barcelona answered “No” and Madrid reported it as “In progress”. It should be noted that, for Valencia, the value displayed corresponds to the most senior official, since the three respondents did not fully coincide on the question of internal resistance.

Table 3. Categorical responses to the closed interview questions, Governance dimension, by city.

Dimension	Barcelona	Madrid	Valencia	Valladolid
Internal resistance to integration	No	Unknown	Yes	Yes
Exchange mechanism with other municipalities	Yes	In progress	Yes	Yes
Exchange mechanism, international scope	No	In progress	Yes	Yes

4.3. Quantitative Results: Cross-City Comparison of Likert Assessments with R

The quantitative analysis, conducted using the statistical software R (The R Foundation for Statistical Computing, Vienna, Austria; version 4.5.2), yielded the following results. As these scores reflect the self-assessment of the officials directly responsible for each urban sandbox rather than an independent evaluation, they are best read as an indicator of self-perceived performance. Section 5.3 cross-checks the most convergent of these results against the qualitative evidence gathered, to assess the extent to which social desirability bias may be inflating them. Figure 4 reports the mean Likert scores (1–6) for each city, with error bars showing ± 1 SD. Valencia’s score is computed as the mean of its three respondents. Barcelona records the highest self-reported mean (5.29), followed by Madrid and Valencia, which coincide exactly at 4.86. Valladolid reports the lowest mean

(4.57). The four means are concentrated within a narrow band of less than one scale point. The error bars, however, behave differently across cities: they remain comparatively narrow for Madrid and Valladolid, whereas for Valencia the ± 1 SD interval spans from 3.0 to 6.0, the widest range of the four cities, and for Barcelona the interval spans from 4.0 to 6.0.

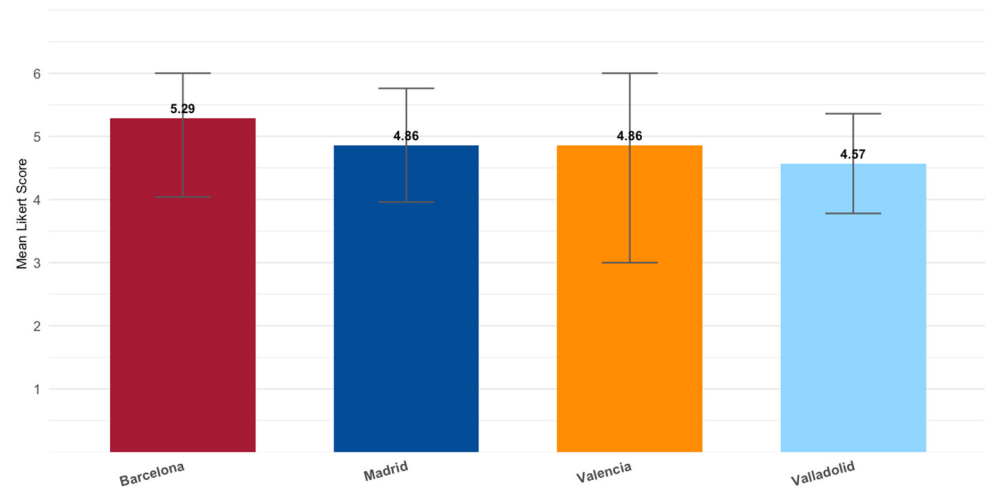


Figure 4. Mean self-reported Likert score by city.

Beyond the city means, the boxplot in Figure 5 displays the underlying distribution of the scores, with each dot corresponding to a city-level mean per question. The two boxes span from 4.0 to 5.0 and from 5.0 to 6.0 respectively, together covering the upper half of the scale, where the majority of points are concentrated. Three points fall outside these boxes, below the 4.0 threshold: one from Barcelona (around 3.0), one from Valladolid (around 2.8), and one from Valencia (around 1.4), the latter constituting the lowest value recorded. Madrid is the only city with no points below 3.9. Within the upper box (5.0–6.0), Barcelona and Madrid each contribute two points, while Valencia contributes a single point. Within the lower box (4.0–5.0), Valladolid contributes the largest share, with four of its seven values falling within that range.

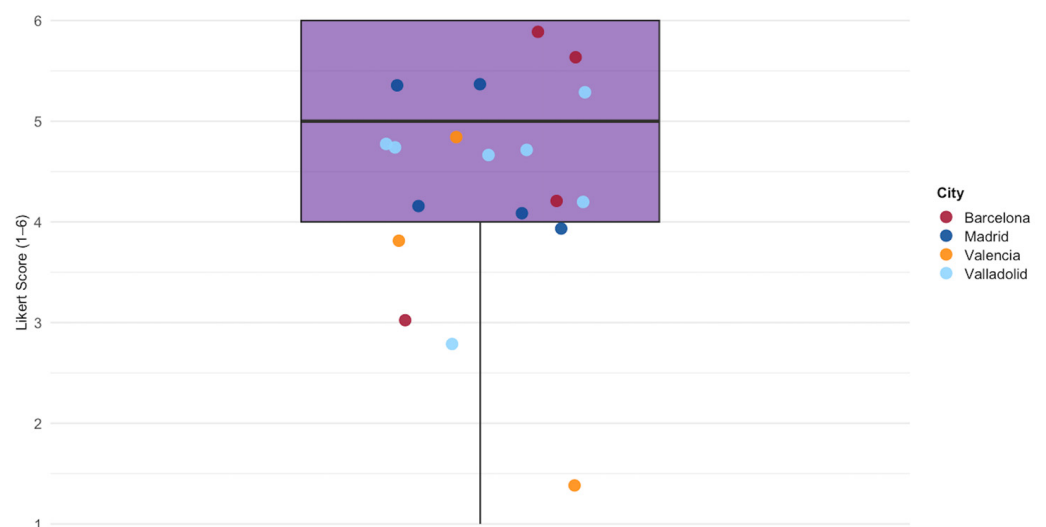


Figure 5. Cross-city patterns in the Likert assessment of the Governance dimension. The vertical line beneath the boxes corresponds to the lower whisker, extending to the lowest recorded value in the distribution.

Additionally, an analysis in spider charts was conducted. Figure 6 compares the four cities across the three dimensions (scale 1–6) of this study. The four urban sandboxes converge most closely on flexibility vs. rigidity, where the lines from all cities sit within a narrow band near the outer rings. The clearest divergence appears on the Learning vs. Isolation axis, where Barcelona reaches the maximum value while the remaining three cities sit visibly further in. On the Governance & Climate Mission axis, the ordering of cities differs from the other two dimensions: Valladolid extends furthest, ahead of both Barcelona and Madrid, a position it does not hold on either of the other axes.

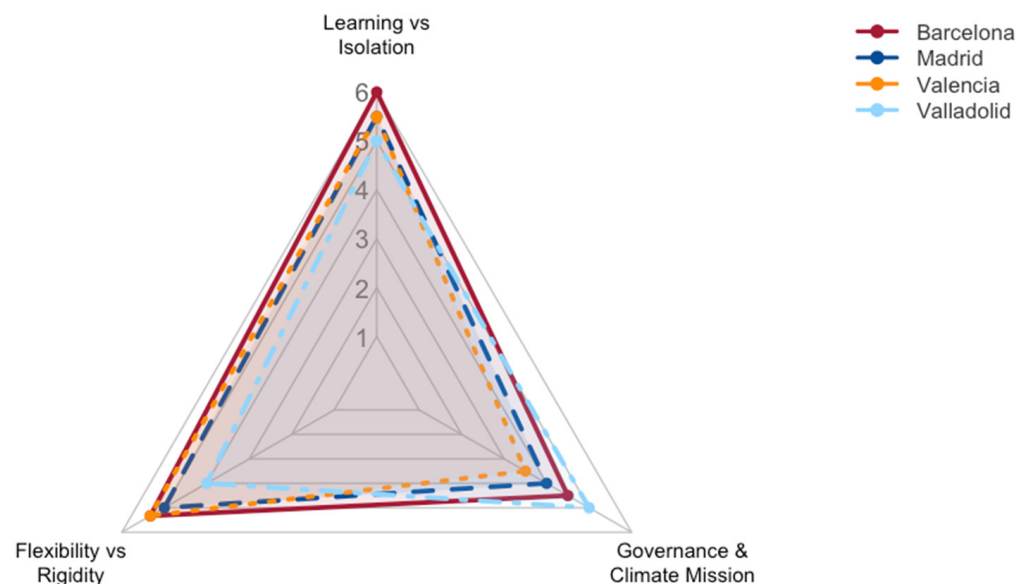


Figure 6. Mean self-reported Likert scores by dimension across cities.

Perceptions related to each of the governance dimensions showed considerable variation, with values spanning from 1.0 to 6.0 (see Figure 7). Within institutional learning vs. isolation, systematisation ranged from 5.0 in Madrid, Valencia and Valladolid to 6.0 in Barcelona (mean 5.2), while interdepartmental collaboration scored 6.0 in every city except Valladolid, which recorded 5.0 (mean 5.8). The flexibility vs. bureaucratic rigidity dimension showed mixed patterns across its three items. Response time ranged from 3.0 in Valladolid to 6.0 in Barcelona and Valencia (mean 5.0); regulatory flexibility ranged from 4.0 in Valencia to 6.0 in Barcelona and Madrid (mean 5.2); and risk tolerance followed an inverse pattern, with Valencia at 6.0 against 4.0 in the remaining three cities (mean 4.5). The sharpest contrast emerged within governance & climate mission: multi-level coordination scored 1.0 in Valencia, the lowest single value recorded across the dataset, against 3.0 in Barcelona, 4.0 in Madrid and 5.0 in Valladolid (mean 3.2); external stakeholder participation, by contrast, returned high and comparatively even scores, at 6.0 in Barcelona and Valencia, 5.0 in Valladolid and 4.0 in Madrid (mean 5.2). Overall, Barcelona and Madrid tended to score at the upper end of the scale, whereas Valencia combined the highest single value (risk tolerance, 6.0) with the lowest (multi-level coordination, 1.0).

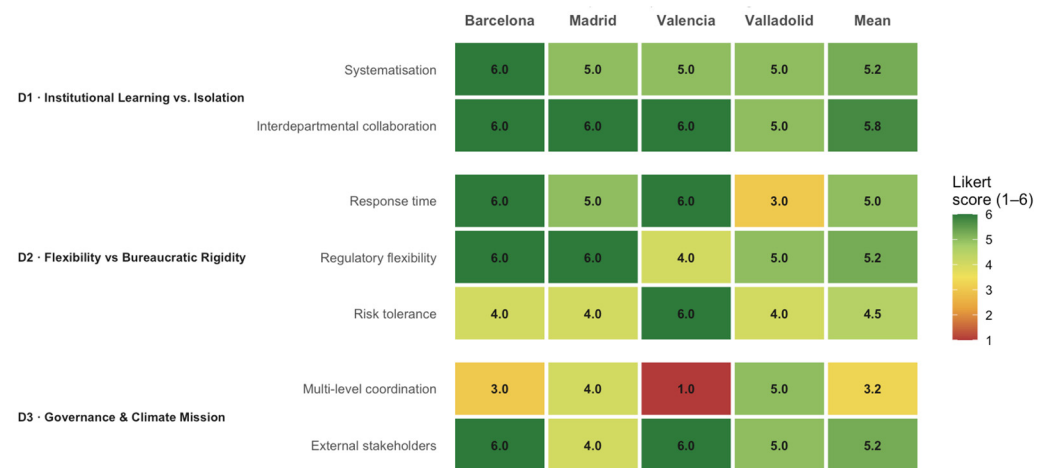


Figure 7. Per-question self-reported Likert scores by city. Bold values are used for legibility against the coloured background and do not encode additional information; colour intensity reflects the Likert score (1–6), as shown in the accompanying scale.

5. Discussion

5.1. Governance Capacities and Institutional Change: An Asymmetric Interaction

The tension between climate urgency and municipal administrative principles that motivated this study (Section 1) is precisely what the evidence gathered here reflects. Across the three governance dimensions assessed, learning vs. isolation, flexibility vs. rigidity and climate mission alignment, the qualitative and quantitative evidence merges on a single pattern: urban regulatory sandboxes in the Spanish Mission for Climate-Neutral and Smart Cities generate real governance capacities, but these capacities remain only partially, and unequally, institutionalised. Cities report strong knowledge circulation, express high self-perceived flexibility, and mobilise external stakeholders effectively, the kind of inclusion that mission-oriented innovation policy explicitly seeks to encourage [46,47]; yet the same evidence shows that this knowledge is not easily internalised, that flexibility is more easily claimed than exercised, and that vertical coordination remains structurally constrained regardless of an urban sandbox's internal performance. The sections that follow interpret this asymmetry as a set of tensions, barriers, and opportunities for institutionalising these capacities, rather than as mere differences between cases: a tension between the speed of organisational learning and its organisational absorption (Section 5.2), an institutional and jurisdictional barrier bounding how far regulatory flexibility can travel (Section 5.3), and, in the paradox this pattern ultimately reveals, what institutionalisation would require of municipal bureaucracy going forward (Section 5.4).

5.2. Knowledge Circulation and Its Limits: Learning as the More Decisive Driver of Adaptive Governance

The clearest evidence from this study comes from the Learning vs. Isolation dimension. Systematisation and interdepartmental collaboration were among the highest-scoring items in the dataset, and the semantic network built from the interviews places knowledge sharing as the single most connected node. This is consistent with the integrated, cross-departmental experimentation the literature treats as a precondition for moving beyond reactive climate policy [34]. Yet the same network identifies only one contradictory relation in the entire Governance code group, linking interdepartmental collaboration directly to internal resistance to integration. The categorical data corroborate this at a broader level: city councils were unanimous on almost every other closed question, but split entirely on internal resistance, with Valencia and Valladolid reporting it, Barcelona denying it, and Madrid describing it as unresolved.

This is precisely the gap between adaptive-governance ambition and adaptive-governance practice that Boyd and Juhola [32] document more broadly, and it speaks to the “pilot paradox”, introduced by van Buuren et al. [11] and examined in depth by Gartlinger and Gualini [53]: on the internal-success side, sandboxes can build formal mechanisms for sharing knowledge—the semantic network’s own most connected node—but on the external, systemic side, whether that knowledge becomes institutionalised, or simply accumulates alongside the routines it was meant to inform, remains open and city-specific. That knowledge sharing is simultaneously the most connected node in the network and the node implicated in its only contradiction suggests that learning infrastructure has advanced beyond the organisational resolution of resistance to using it. The speed at which knowledge circulates appears to have been gained partly at the expense of the slower, more contested work of embedding it into standing municipal routines, a pattern consistent with the practice-level account of institutional change proposed by Alméstar and Romero-Muñoz [35]. The Valladolid interviewee illustrated this directly, describing resistance as primarily technical-operational: “staff is aware of day-to-day service delivery and find it harder to recognise the medium-term benefits of an innovative project than its immediate operational inconveniences” (Valladolid City Hall, interview).

5.3. Flexibility Within Limits: Jurisdiction, Discretion, and the Price of Legal Certainty

Flexibility vs. Rigidity is the dimension on which the four cities converge most closely. Reading against the distinction between formal and substantive flexibility [9,29], this convergence is as much a methodological flag as a substantive finding. Self-reported flexibility, scored by the practitioners responsible for operating the instrument, a textbook setting for social desirability bias, is exactly the indicator on which a formally adaptive but substantively unchanged arrangement would be expected to score high, since perceiving flexibility is easier and faster to achieve than redistributing the authority or resources that would make it durable [41]. A comparable pattern is reported by Kim and Cho [49], who find that in Korea’s regulatory sandbox scheme, government policy orientation systematically moderates how quickly sandbox-generated learning translates into binding regulatory reform, evidencing that the gap between perceived adaptiveness and enacted change is not unique to the Spanish case. Two pieces of evidence support this perspective. First, some individual items diverge considerably: response time ranges from 3.0 in Valladolid to 6.0 in Barcelona and Valencia, and risk tolerance follows an inverted pattern in which Valencia alone scores 6.0 against 4.0 elsewhere (Figure 7), echoing Hooghe and Marks’ [25] observation that municipal climate governance sits between the durable logic of Type I governance and the bounded, task-specific logic of Type II. Second, Climate Mission Alignment, framed around specific coordination mechanisms rather than general self-evaluation, produced the most divergent and single lowest score in the entire dataset (Valencia, multi-level coordination, 1.0/6.0), consistent with perceived flexibility being easier to claim than multi-level coordination is to perform, and consistent with the translation cost that mission-oriented innovation frameworks acquire once they meet the resource-constrained reality of city government [48].

The interviews clarify why this limitation is jurisdictional as much as organisational. In Valencia, vertical coordination with other tiers of government “doesn’t really exist”: “national science legislation governing testing spaces explicitly excludes local administrations from competence in this area” (Valencia City Hall, interview), locating the constraint in the surrounding multi-level architecture rather than in the urban sandbox instrument itself. Madrid and Barcelona illustrate the same boundary from within municipal competence: the Madrid respondent described the urban sandbox as granting “regulatory exemptions, whatever is necessary for trials such as street-operating robots” (Madrid City

Hall, interview), whereas Barcelona’s account draws a sharper line: “officials can adjust thresholds within their own competence, for instance, drone-flight altitude, but cannot exempt regulation belonging to another administration, at most pursuing dialogue where a request touches a state-level competence such as energy-grid integration” (Barcelona City Hall, interview). The Valladolid case makes the underlying trade-off explicit: an ordinance, one official noted, “gives legal certainty, but is at the same time rigid, since it is drafted with a particular direction in mind and cannot anticipate every situation that later arises” (Valladolid City Hall, interview). Flexibility, in summary, is real within a municipality’s own jurisdiction, where it converts into negotiation rather than exemption.

5.4. Relocating Bureaucratic Discretion: The Paradox at the Heart of Urban Sandboxes

Taken together, and despite all four cities operating under broadly the same national legal basis [12] and the same supranational mission [7], the city-level profiles suggest that sandboxes do not remove bureaucratic mediation from municipal governance; they reconfigure it. To interpret this asymmetry, it is useful to draw on Scott’s [54] distinction between three pillars through which institutions are carried and sustained: a regulative pillar of formal rules and sanctioning mechanisms, a normative pillar of professional norms and standard operating procedures, and a cultural-cognitive pillar of taken-for-granted assumptions and shared understandings. Read through this lens, the formal ordinances that create each urban sandbox (the regulative pillar) remain essentially unchanged from their initial design, while the four city profiles below diverge chiefly in how far cross-departmental routines (the normative pillar) and the taken-for-granted practices of individual facilitators (the cultural-cognitive pillar) have absorbed the knowledge each sandbox generates. Barcelona records the highest overall mean and reaches the ceiling on Learning vs. Isolation, yet does not stand out on multi-level coordination: a profile of strong internal learning loosely coupled to the vertical governance architecture above it. Madrid shows the most consistently mid-to-high scores, with no item below 3.9, consistent with institutional flexibility achieved through the practised reinterpretation of existing rules documented for Madrid’s collaborative planning processes [35] rather than through structural redesign, a redistribution of practical authority without formal reform that echoes comparable governance transitions documented elsewhere, such as the shift toward resident-led spatial self-management in Chinese urban vending zones [36]. Valladolid combines the lowest overall mean with a decoupled mission-alignment result: it scores near the bottom on the other two dimensions yet extends furthest of the four cities on Climate Mission Alignment, having formalised a CCC and aligned its public discourse with the Mission’s targets [55] without this necessarily translating into faster learning or more flexible internal operation. This decoupling is consistent with the limited involvement of implementing actors in shaping the meta-governance framework itself that Gartlinger and Gualini [53] document in the Berlin case.

Valencia makes the reconfiguration most explicit. Rather than amending the underlying regulatory framework, exemptions are negotiated case by case through an internal procedure that attaches conditions to each trial: a robot-delivery pilot near Valencia’s cathedral, for example, was authorised on condition that the robot stay ten metres clear of the building, a markedly more permissive condition than other applicants in the same area would receive (Valencia City Hall, interview), gaining adaptability at the price of uniform treatment across applicants. This case-by-case, person and project-dependent route to flexibility is consistent with Valencia’s own internal inconsistency, the highest score in the dataset (risk tolerance) and the lowest (multi-level coordination) both belong to it, and with the fact that its three respondents did not fully agree on whether internal resistance to integration exists. The urban sandbox, in this sense, does not dissolve bureaucratic

discretion; it relocates it from the general rule to the individual case, replacing ex ante regulatory certainty with ex-post negotiated permission.

This is the governance paradox this study identifies: the same instrument that generates measurable learning, perceived flexibility, and stakeholder engagement does so largely by redistributing bureaucratic judgment to individual facilitators and case-by-case procedures, a shift concentrated in the normative and cultural-cognitive pillars of institutionalisation, in Scott's [54] terms, rather than by transforming the regulative structures that bureaucracy sits within. This pattern also speaks to Park's [41] concept of Dynamic Regulatory Learning: a cyclical process in which pilot implementation generates empirical feedback that is fed back into policy redesign, in turn reshaping both governance and market behaviour. Read against this model, the Spanish cases occupy different points of the same loop: in cities where cross-departmental routines already channel sandbox-generated knowledge into standing practice, feedback closes back into policy design in the way Park's model anticipates; in cities where this knowledge remains dependent on individual facilitators, the loop stalls before reaching policy redesign—precisely the disconnection the pilot paradox [11] describes. Park's comparative benchmarking of sandbox regimes in Korea, Singapore, and China [41] further underscores what is distinctive about the Spanish, EU-embedded case examined here: unlike Korea's centrally coordinated, mission-oriented model, Spanish sandboxes operate within a multi-level governance architecture in which municipal, regional, and national competences are only loosely connected—itself a plausible explanation for why the loop closes unevenly across the four cities. Read together with Okonjo's [45] argument that public sector sandboxes require substantial institutional adaptation before they can generate legitimate, durable outcomes, this suggests that the normative and cultural-cognitive work documented in this study is what determines whether Dynamic Regulatory Learning, in Park's sense, actually closes. These capacities are not merely an internal administrative concern: Valencia's own account of sandbox operation as "projecting the startup ecosystem" abroad through international missions and delegations while "bringing in" startups from other cities to trial their solutions locally (Valencia City Hall, interview) suggests they condition, too, how legible and attractive the surrounding entrepreneurial ecosystem becomes to the firms and innovators sandboxes are meant to serve. Whether this redistributed, person-dependent flexibility is a durable substitute for institutional change, or simply a slower route to it, remains the central open question for future research on adaptive governance in urban sandboxes.

6. Conclusions

6.1. Synthesis of Findings

This study examined the conditions under which urban regulatory sandboxes can foster adaptive governance capacities in cities, taking as its case study four Spanish municipalities participating in the Mission for Climate-Neutral and Smart Cities. The evidence indicates that the answer is neither a simple yes nor a simple no. Urban sandboxes are perceived by the practitioners interviewed to foster adaptive governance logics, but the depth of that adaptation depends heavily on which dimension of governance is examined, and considerably less on which city is examined, since all four municipalities operate under a broadly similar national legal basis and the same supranational mission. Therefore, the cross-sectional design of this study cannot rule out that pre-existing differences in municipal governance capacity also contribute to this pattern (see Section 6.3).

The most robust finding concerns learning vs. isolation. Within this, formal mechanisms for knowledge circulation are broadly in place across all four sandboxes, but whether this knowledge becomes embedded in standing departmental practice, escaping the "pilot paradox" that the urban experimentation literature repeatedly documents [11,53], remains

on an internal resistance that varies markedly from one city to the next. This is, at the same time, the most actionable lever identified in this study. Unlike multi-level coordination, which depends on actors and competences beyond the municipality's direct control, internal resistance to integration is a governance variable that municipal leadership can address directly, for instance through dedicated cross-departmental routines for absorbing sandbox-generated knowledge, rather than relying on the existence of a knowledge-sharing platform alone.

Climate mission alignment emerged as the weakest and most polarised dimension, and the one most exposed to the friction between supranational mission, national legal framework, and municipal practice that the theoretical framework anticipated [48]. An urban sandbox cannot, by itself, manufacture the vertical coordination that this alignment requires. Where it is weak, as in Valencia's near-floor score on multi-level coordination, the limitation appears to lie above the urban sandbox, in the surrounding multi-level governance architecture, rather than within the instrument's own design. Alternatively, the partial disconnection of mission alignment from internal flexibility observed in Valladolid suggests that signing a CCC and operating an adaptively governed urban sandbox are related but distinct achievements, and that policymakers should resist treating formal mission participation as a proxy for substantive local adaptive capacity.

The convergence observed on flexibility vs. rigidity should be read with particular caution. This dimension was assessed through the self-reports of the practitioners who run the instrument, and it diverged sharply from more concrete items nested within it, such as response time. For this reason, the high and converging scores recorded here are better read as evidence of formal, perceived adaptiveness than as confirmation of real institutional change [29,41].

For municipal policymakers, these findings carry a clear practical implication: urban regulatory sandboxes are a necessary but not sufficient condition for embedding governance adaptability within municipal governance itself.

6.2. Policy Implications: Municipal Action Versus Multi-Level Negotiation

What appears to be the differentiator, based on the evidence gathered in this study, is the deliberate institutional work that surrounds the urban sandbox, work that falls into two distinct categories. The first category lies within a municipality's own competence. Transforming sandbox-generated knowledge into standing practice, rather than leaving it dependent on individual facilitators, is achievable through dedicated cross-departmental routines: the internal resistance to integration documented in Section 5.2, and the case-by-case, person-dependent route to flexibility documented in Section 5.4, are governance variables a city government can address directly, by embedding what currently operates as informal, cultural-cognitive practice into the normative pillar of standing procedure [43], for example, formalising the follow-up mechanisms that currently depend on which official happens to run the sandbox. The second category depends on actors and competences beyond the municipality's direct control. Climate Mission Alignment, the weakest and most polarised dimension in this study, is constrained less by any single sandbox's design than by the multi-level governance architecture surrounding it [25]. Valencia's account of national science legislation excluding local administrations from competence over testing spaces illustrates a gap that only reform above the municipal level can close whether through the national regulatory framework established by Royal Decree 568/2022 [12] or through the Mission's own coordination structures [7]. For municipal policymakers, the actionable driver here is not implementation but advocacy: pressing regional and national authorities for standing liaison mechanisms between local sandboxes and mission-level

governance, rather than relying on each city to negotiate this connection informally and case by case, as currently happens.

6.3. Limitations and Future Research

Future research should test this proposition, regarding the convergence observed on flexibility vs. rigidity, directly by triangulating perceived flexibility against harder administrative indicators, such as actual processing times, the number and nature of regulatory exemptions granted, or the rate at which sandbox outcomes are translated into permanent regulatory amendments.

A longitudinal design, ideally paired with a larger sample capable of tracking urban sandboxes as they mature beyond their initial implementation phase, would help establish whether the municipal-level and multi-level patterns identified in this study strengthen, persist, or dissolve as Spain's urban sandbox ecosystem develops further. Doing so would also situate Spain's experience within the broader trajectory of urban resilience research, which literature evidence suggests is moving toward digitalisation, artificial intelligence, and multi-scale collaborative governance as its next frontier [33].

These interpretations are bound by the same limitations that shaped the analytical strategy. With six interviews distributed across four cities, and three of them concentrated in a single municipality, the dataset does not support inferential statistical testing, and all comparative claims made above are descriptive rather than confirmatory. Similarly, all four cases involve an active sandbox; the design includes no discontinued or unsuccessful case, and so cannot isolate what separates success from failure in the stronger sense this would require. As noted in Section 3.2, these constraints reflect the still-small population of officials with direct operational responsibility for an active urban sandbox, given how recently most Spanish municipalities have adopted the instrument. The cross-sectional design, dictated by the recent diffusion of urban sandboxes across Spanish municipalities between 2022 and 2026, also means that the present analysis captures an early stage of implementation rather than a mature state; some of the divergence observed, particularly on flexibility vs. rigidity and climate mission alignment, may narrow or widen as sandboxes accumulate operating experience. This is most visible in Barcelona's case: at the time of interview, its urban sandbox had only recently been formally approved, so the practitioner's assessment likely reflects the quality of the newly adopted regulatory design as much as lived operational experience, and should be read with corresponding caution. This study is further limited to the perspective of the officials who run each sandbox: the extent to which residents, startups, or civil-society organisations perceive their needs as met by the urban sandbox, and what those needs are in their own terms, is not addressed here. Systematically comparing practitioners' self-assessments against this beneficiary-side perspective is an important avenue for future research. Two further steps could strengthen this evidentiary base: a documentary review of municipal climate plans, urban sandbox meeting minutes, and evaluation reports, and a desktop framework benchmarking each city's sandbox ecosystem on dimensions such as mayoral support, administrative autonomy, funding stability, and EU-funded projects. More fundamentally, the three governance dimensions examined here are assessed through practitioners' own accounts of their urban sandbox's operation, not against independent institutional indicators such as formal reforms, budgetary records, meeting minutes, or external evaluation reports; incorporating such indicators, where they become available as sandboxes mature, is an important direction for future research. Such indicators should also cover a more fine-grained set of capacity-related variables, including decision-making and evidence use, management coordination, regulatory output, political legitimacy, and built-form project delivery. Finally, the Likert-scale responses came from practitioners directly responsible for the instrument being evaluated, so social-

desirability bias cannot be excluded; these same officials also hold a vested interest in their sandbox's perceived success, which may reinforce rather than offset this bias. This is why the convergence observed on self-reported flexibility is treated above as a methodological flag rather than as decisive evidence of real adaptive capacity. More broadly, because these conclusions are drawn from a single case study situated in one national and supranational governance context, extending this analysis to urban regulatory sandboxes operating under different geographic and institutional conditions remains an important avenue for future research.

Author Contributions: Conceptualization, A.H.-B., M.A. and E.N.-C.; methodology, A.H.-B.; software, A.H.-B.; validation, A.H.-B., M.A. and E.N.-C.; formal analysis, A.H.-B.; investigation, A.H.-B.; resources, A.H.-B.; data curation, A.H.-B.; writing—original draft preparation, A.H.-B.; writing—review and editing, A.H.-B., M.A., C.A. and C.F.; visualization, A.H.-B., M.A. and E.N.-C.; supervision, A.H.-B., M.A., E.N.-C., C.A. and C.F.; project administration, A.H.-B. and E.N.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This study received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study by the Ethics Committee for R&D+i Activities at the Universidad Politécnica de Madrid as per Article N°1, section 1.1 of the Regulations of the Ethics Committee for R&D+i Activities at the Universidad Politécnica de Madrid (approved by the Governing Board on 30 March 2017), submission to the Committee is required only for research involving human subjects or human samples, the collection or processing of personal data affecting fundamental rights, or activities such as animal testing, biological/chemical agents, or radioactive substances.

Informed Consent Statement: Verbal informed consent was obtained from the participants. Verbal consent was obtained rather than written because all participants were informed of the purpose of the research, the intended use of their responses, the anonymized and academic nature of the data collected, and their right to withdraw at any time.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy and confidentiality restrictions, as the interview data contain identifiable statements from municipal practitioners who participated under conditions of confidentiality.

Acknowledgments: The authors thank the municipal officials of Barcelona, Madrid, Valencia, and Valladolid City Councils for their participation in the interviews and their support in facilitating access to relevant documentation. During the preparation of this manuscript, the authors used Claude (Anthropic; Sonnet 4.6) for language and structural editing, and for assistance with the R code used. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CCC Climate City Contract
EU European Union

Appendix A

Semi-structured interview model

General Information

Identifying details of the interviewee, collected at the start of the interview:

- First name
- Last name

- Position/Role
- Unit
- Service/Department
- Deputy Directorate-General
- Government Area
- City Council represented

Table A1. Semi-structured interview protocol—Dimension 1: Learning vs. Isolation.

Question	Type	How It Is Measured
To what extent are the sandbox's results being systematised to improve the City Council's regulatory frameworks or, alternatively, is there a defined protocol for them to serve as the basis for future reforms in municipal procedures and policies?	Likert	1–6 scale (ordinal/interval)
Have there been internal, cultural, legal, or political resistances to integrating the sandbox's results into municipal operations?	Open	Qualitative response
If so, what type of internal resistance has occurred?	Open	Qualitative response
To what extent has the sandbox fostered interdepartmental collaboration, whether through the involvement of different areas in its design and implementation, or through (formal or informal) mechanisms established to transfer learning to other municipal departments?	Likert	1–6 scale (ordinal/interval)
Is there a body or figure responsible for systematising the sandbox's learnings and, beyond internal transfer, have channels been established to share this knowledge with other municipalities (whether within the same autonomous community or with city councils in other regions)?	Open	Qualitative response

Table A2. Semi-structured interview protocol—Dimension 2: Flexibility vs. Bureaucratic Rigidity.

Question	Type	How It Is Measured
To what extent do you consider that the sandbox reduces administrative response times compared to ordinary procedures?	Likert	1–6 scale (ordinal/interval)
To what extent do you consider that the sandbox has mechanisms that allow existing regulations to be adapted or made more flexible to facilitate the development of pilot projects?	Likert	1–6 scale (ordinal/interval)
What is the specific procedure when a project requires an exception or adaptation of current regulations?	Open	Qualitative response
Following on from the previous question, who has the authority to authorise it?	Open	Qualitative response
Do you consider that the sandbox has fostered a culture of greater risk tolerance within the administration?	Likert	1–6 scale (ordinal/interval)

Table A3. Semi-structured interview protocol—Dimension 3: Governance and Climate Mission.

Question	Type	How It Is Measured
To what extent do you consider that there is effective coordination with other levels of government (regional, national, or European) to facilitate the scaling-up or integration of the solutions developed in the sandbox?	Likert	1–6 scale (ordinal/interval)
To what extent do you consider that external actors (companies, startups, universities, citizens) participate in defining the sandbox's objectives, criteria, and decisions?	Likert	1–6 scale (ordinal/interval)

Table A3. Cont.

Question	Type	How It Is Measured
Is there any exchange mechanism with other municipalities that also operate urban sandboxes at the national level?	Closed	Dichotomous (Yes/No)
Following on from the previous question, if so, could you provide more information?	Open	Qualitative response
And in the international case?	Closed	Dichotomous (Yes/No)
If so, could you elaborate further?	Open	Qualitative response

References

- Bulkeley, H.; Betsill, M. *Cities and Climate Change: Urban Sustainability and Global Environmental Governance*; Routledge: London, UK, 2003. [\[CrossRef\]](#)
- Bulkeley, H. Cities and the Governing of Climate Change. *Annu. Rev. Environ. Resour.* **2010**, *35*, 229–253. [\[CrossRef\]](#)
- Castán Broto, V. Climate change politics and the urban contexts of messy governmentalities. *Territ. Polit. Gov.* **2020**, *8*, 241–258. [\[CrossRef\]](#)
- Alméstær Urteaga, J.M. La Ciudad Como Prototipo: Estrategias, Herramientas y Aprendizajes de la Experimentación Urbana Ante la Crisis Ecológica. Ph.D. Thesis, Universidad Politécnica de Madrid, Madrid, Spain, 2025. [\[CrossRef\]](#)
- Hildén, M.; Jordan, A.; Huitema, D. Special Issue on Experimentation for Climate Change Solutions Editorial: The search for climate change and sustainability solutions: The promise and the pitfalls of experimentation. *J. Clean. Prod.* **2017**, *169*, 1–7. [\[CrossRef\]](#)
- Alaassar, A.; Mention, A.-L.; Aas, T.H. Exploring a new incubation model for FinTechs: Regulatory sandboxes. *Technovation* **2021**, *103*, 102237. [\[CrossRef\]](#)
- European Commission; Joint Research Centre. *Regulatory Sandboxes for Net-Zero Innovation—What Energy Sector Experimentation Reveals for Policy and Regulation*; Publications Office of the European Union: Luxembourg, 2026. Available online: <https://publications.jrc.ec.europa.eu/repository/handle/JRC145146> (accessed on 11 July 2026).
- Torrens, J.; Schot, J. The Roles of Experimentation in Transformative Innovation Policy. 2017. Available online: <https://tipconsortium.net/wp-content/uploads/2018/03/Research-Brief-for-Experimentation.pdf> (accessed on 12 July 2026).
- Wolfram, M. Conceptualizing urban transformative capacity: A framework for research and policy. *Cities* **2016**, *51*, 121–130. [\[CrossRef\]](#)
- Evans, J.; Karvonen, A.; Raven, R. The experimental city: New modes and prospects of urban transformation. In *The Experimental City*; Routledge: London, UK, 2016; pp. 1–12. [\[CrossRef\]](#)
- van Buuren, A.; Vreugdenhil, H.; van Popering-Verkerk, J.; Ellen, G.J.; van Leeuwen, C.; Breman, B. The Pilot Paradox: Exploring Tensions between Internal and External Success Factors in Dutch Climate Adaptation Projects. In *Innovating Climate Governance: Moving Beyond Experiments*; Turnheim, B., Kivimaa, P., Berkhout, F., Eds.; Cambridge University Press: Cambridge, UK, 2018; pp. 145–165. Available online: <https://www.cambridge.org/core/books/abs/innovating-climate-governance/pilot-paradox/0E34E28C2E7E44BB452AC2A88CE596D1> (accessed on 15 May 2026).
- Real Decreto 568/2022, De 11 De Julio, Por El Que Se Establece El Marco General Del Banco De Pruebas Regulatorio Para El Fomento De La Investigación Y La Innovación En El Sector Eléctrico. Boletín Oficial Del Estado, 12 July 2022, No. 166. Available online: <https://www.boe.es/eli/es/rd/2022/07/11/568/con> (accessed on 17 May 2026).
- Hall, P.A. Policy paradigms, social learning, and the state: The case of economic policymaking in Britain. *Comp. Polit.* **1993**, *25*, 275–296. [\[CrossRef\]](#)
- Borins, S. Encouraging innovation in the public sector. *J. Intellect. Cap.* **2001**, *2*, 310–319. [\[CrossRef\]](#)
- Williams, P. The competent boundary spanner. *Public Adm.* **2002**, *80*, 103–124. [\[CrossRef\]](#)
- Lafferty, W.M.; Hovden, E. Environmental policy integration: Towards an analytical framework. *Environ. Polit.* **2003**, *12*, 1–22. [\[CrossRef\]](#)
- Hartley, J. Innovation in governance and public services: Past and present. *Public Money Manag.* **2005**, *25*, 27–34. [\[CrossRef\]](#)
- Organización para la Cooperación y el Desarrollo Económicos (OCDE). *Fostering Innovation in the Public Sector*; OECD Publishing: Paris, France, 2017. [\[CrossRef\]](#)
- Kattel, R.; Mazzucato, M. Mission-oriented innovation policy and dynamic capabilities in the public sector. *Ind. Corp. Change* **2018**, *27*, 787–801. [\[CrossRef\]](#)
- Omarova, S.T. Technology v technocracy: Fintech as a regulatory challenge. *J. Financ. Regul.* **2020**, *6*, 75–124. [\[CrossRef\]](#)

21. Cauffman, C.; Goanta, C. A new order: The Digital Services Act and consumer protection. *Eur. J. Risk Regul.* **2021**, *12*, 758–774. [CrossRef]
22. Hintz, M.J.; Kaack, L.H.; Creutzig, F.; Vitale, T. Framing artificial intelligence as a policy instrument in urban climate action: Practitioners' perspectives from Paris. *SSRN Electron. J.* **2026**. advance online publication. [CrossRef]
23. Rose, R. What is Lesson-Drawing? *J. Public Policy* **1991**, *11*, 3–30. [CrossRef]
24. Dolowitz, D.P.; Marsh, D. Learning from abroad: The role of policy transfer in contemporary policy-making. *Governance* **2000**, *13*, 5–23. [CrossRef]
25. Hooghe, L.; Marks, G. Unraveling the Central State, But How? Types of Multi-Level Governance. *Am. Polit. Sci. Rev.* **2003**, *97*, 233–243. [CrossRef]
26. Casado-Asensio, J.; Steurer, R. Integrated strategies on sustainable development, climate change mitigation and adaptation in Western Europe: Communication rather than coordination. *J. Public Policy* **2014**, *34*, 437–473. [CrossRef]
27. Karvonen, A. The City of permanent experiments? In *Innovating Climate Governance*; Turnheim, B., Kivimaa, P., Berkhout, F., Eds.; Cambridge University Press: Cambridge, UK, 2018; pp. 201–215. [CrossRef]
28. Goldsmith, S.; Eggers, W.D. *Governing by Network: The New Shape of the Public Sector*; Brookings Institution Press: Washington, DC, USA, 2004. Available online: <http://www.jstor.org/stable/10.7864/j.ctt12879qp> (accessed on 9 May 2026).
29. Moyo, T.N.; Paramasivam, C.; Mashingaidze, T.M. Multi-level governance and climate change adaptation: A systematic literature review. *Sustain. Futures* **2026**, *11*, 101724. [CrossRef]
30. Dietz, T.; Ostrom, E.; Stern, P.C. The struggle to govern the commons. *Science* **2003**, *302*, 1907–1912. [CrossRef] [PubMed]
31. Folke, C.; Hahn, T.; Olsson, P.; Norberg, J. Adaptive governance of social-ecological systems. *Annu. Rev. Environ. Resour.* **2005**, *30*, 441–473. [CrossRef]
32. Boyd, E.; Juhola, S. Adaptive climate change governance for urban resilience. *Urban Stud.* **2015**, *52*, 1234–1264. [CrossRef]
33. Han, M.; Fu, G.; Wu, Z.; Lu, Y.; Xie, X.; Xu, S. A systematic review of the trajectory of urban resilience research: A bibliometric perspective on global trends and China's pathway. *Sustainability* **2026**, *18*, 2945. [CrossRef]
34. Anguelovski, I.; Carmin, J. Something Borrowed, Something New: Innovation and Institutionalization in Urban Climate Adaptation. *Urban Stud.* **2011**, *48*, 1697–1715. [CrossRef]
35. Alméstar, M.; Romero-Muñoz, S.; Mestre, N. Breaking Silos: A Systemic Portfolio Approach and Digital Tool for Collaborative Urban Decarbonisation. *Sustainability* **2025**, *17*, 5145. [CrossRef]
36. Zhai, Y.; Ma, P.; Li, M. Sustainable development and environmental governance for urban vending zones: A case study in the Waliu Community, China. *Sustainability* **2025**, *17*, 4002. [CrossRef]
37. Gupta, J.; Termeer, C.; Klostermann, J.; Meijerink, S.; van den Brink, M.; Jong, P.; Nooteboom, S.; Bergsma, E. The adaptive capacity wheel: A method to assess the inherent characteristics of institutions to enable the adaptive capacity of society. *Environ. Sci. Policy* **2010**, *13*, 459–471. [CrossRef]
38. Beckstedde, E.; Correa Ramírez, M.; Cossent, R.; Vanschoenwinkel, J.; Meeus, L. Regulatory sandboxes: Do they speed up innovation in energy? *Energy Policy* **2023**, *180*, 113656. [CrossRef]
39. Brown, E.; Piroska, D. Governing Fintech and Fintech as Governance: The Regulatory Sandbox, Riskwashing, and Disruptive Social Classification. *N. Polit. Econ.* **2022**, *27*, 19–32. [CrossRef]
40. OECD. Tools for Agility: Actionable Strategic Intelligence and Policy Experimentation. In *OECD Science, Technology and Innovation Outlook 2025*; OECD Publishing: Paris, France, 2025. Available online: https://www.oecd.org/en/publications/oecd-science-technology-and-innovation-outlook-2025_5fe57b90-en/full-report/tools-for-agility-actionable-strategic-intelligence-and-policy-experimentation_288971cb.html (accessed on 9 May 2026).
41. Park, M.J. Dynamic regulatory experimentation and innovation trajectories: Evidence from Korea's battery industry and global sandbox governance. *Technol. Forecast. Soc. Change* **2026**, *226*, 124583. [CrossRef]
42. Attrey, A.; Leshner, M.; Lomax, C. *The Role of Sandboxes in Promoting Flexibility and Innovation in the Digital Age*; OECD Going Digital Toolkit Notes, No. 2; OECD Publishing: Paris, France, 2020. Available online: https://www.oecd.org/en/publications/the-role-of-sandboxes-in-promoting-flexibility-and-innovation-in-the-digital-age_cdf5ed45-en.html (accessed on 17 May 2026).
43. Wady, A.F.; Consoni, F.L. Regulatory sandbox for electric mobility: Enhancing charging infrastructure and innovation. *Renew. Sustain. Energy Rev.* **2025**, *213*, 115496. [CrossRef]
44. Kálmán, J. The Role of Regulatory Sandboxes in FinTech Innovation: A Comparative Case Study of the UK, Singapore, and Hungary. *FinTech* **2025**, *4*, 26. [CrossRef]
45. Okonjo, J. Adapting regulatory sandboxes as experimentalist governance frameworks for public sector artificial intelligence experimentation. *Glob. Public Policy Gov.* **2026**, *6*, 259–281. [CrossRef]
46. Mazzucato, M. *Mission-Oriented Innovation Policy: Challenges and Opportunities*; UCL Institute for Innovation and Public Purpose, Working Paper IIPP WP 2017-1; UCL: London, UK, 2017. Available online: https://discovery.ucl.ac.uk/id/eprint/10195728/1/Mazzucato_moip-challenges-and-opportunities-working-paper-2017-1.pdf (accessed on 9 May 2026).

47. European Commission, Directorate-General for Research and Innovation; Mazzucato, M. *Mission-Oriented Research & Innovation in the European Union: A Problem-Solving Approach to Fuel Innovation-Led Growth*; Publications Office of the European Union: Luxembourg, 2018. [[CrossRef](#)] [[PubMed](#)]
48. Pereverza, K.; Rohrer, H.; Kordas, O. Fostering Urban Climate Transition Through Innovative Governance Coordination. *Environ. Policy Gov.* **2025**, *35*, 631–646. [[CrossRef](#)]
49. Kim, Y.; Cho, K. Socio-technical transitions: Dynamic interactions between actors and regulatory responses in regulatory sandboxes. *Sustainability* **2026**, *18*, 1345. [[CrossRef](#)]
50. Yin, R.K. The case study method as a tool for doing evaluation. *Curr. Sociol.* **1992**, *40*, 121–137. [[CrossRef](#)]
51. Yin, R.K. *Case Study Research and Applications: Design and Methods*, 6th ed.; Sage Publications: Thousand Oaks, CA, USA, 2017. Available online: <https://ebooks.umu.ac.ug/librarian/books-file/Case%20Study%20Research%20and%20Applications.pdf> (accessed on 11 July 2026).
52. Eisenhardt, K.M. Building theories from case study research. *Acad. Manag. Rev.* **1989**, *14*, 532–550. [[CrossRef](#)]
53. Gartlinger, I.; Gualini, E. Climate governance experiments: Current practices and their meta-governance embedding in Berlin's solar energy transition. *Eur. Plan. Stud.* **2025**, *33*, 680–698. [[CrossRef](#)]
54. Scott, W.R. *Institutions and Organizations: Ideas, Interests, and Identities*, 4th ed.; Sage: Thousand Oaks, CA, USA, 2014.
55. Udoviyk, O. A climate-neutral urban Spain: The role of the Mission in advancing the transformative capacity of Spanish cities. *Front. Sustain. Cities* **2025**, *7*, 1694543. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.