

Article

Assessing Public Open Space Quality: Explaining Divergence Across Assessment Tools

Jamal Al-Qawasmi ^{1,2,*} , Md Rafsun Jany ^{1,3} and Ziad Ashour ^{1,4} 

¹ Architecture and City Design Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia; g202426780@kfupm.edu.sa (M.R.J.); zashour@kfupm.edu.sa (Z.A.)

² Interdisciplinary Research Center for Construction and Building Materials, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

³ Department of Architecture, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh

⁴ Interdisciplinary Research Center for Smart Mobility and Logistics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

* Correspondence: jamalq@kfupm.edu.sa

Abstract

Public open space (POS) assessment tools have proliferated across disciplines and application contexts, yet their results often remain difficult to compare. Existing research has largely attributed this divergence to differences in indicator selection, overlooking the broader assessment architecture within which indicators acquire meaning. This study examines why POS assessment tools diverge by conceptualizing them as layered socio-technical sign systems. A systematic literature review identified 38 POS assessment tools, which were analyzed through structured content analysis. Twenty empirically derived attributes were mapped to a five-level semiotic framework comprising the empiric, syntactic, semantic, pragmatic, and social levels. The findings demonstrate that divergence is systematic rather than incidental and extends well beyond indicator content. At the empiric level, 57.9% of tools lack complete operational definitions, while 55.3% rely exclusively on subjective indicators. Structurally, 55.3% function as assessment frameworks rather than fully operational tools, reflecting substantial variation in procedural closure. Across all five semiotic levels, the cumulative interaction of measurement practices, structural organization, conceptual framing, intended application and institutional context generates cross-level structural misalignment that fundamentally constrains comparability. In response, this study proposes a semiotic compliance framework that specifies minimum documentation requirements for improving transparency, internal coherence, and interpretability without prescribing indicator standardization. The findings demonstrate that meaningful comparisons among POS assessment tools depend less on reducing indicator diversity than on explicitly aligning and documenting the multiple semiotic layers through which assessment systems generate meaning and support decision-making. The framework provides practical guidance for developing, evaluating, selecting, and comparing future POS assessment tools while preserving the contextual diversity required across different planning and governance settings.

Keywords: public open space quality; assessment tools; semiotics; semiotic compliance framework; assessment comparability; indicator-based assessment; public space governance



Academic Editor: Jianming Cai

Received: 14 June 2026

Revised: 23 July 2026

Accepted: 24 July 2026

Published: 1 August 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](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Public open spaces (POSs) are publicly accessible outdoor environments that support everyday urban life, including parks, plazas, streets, squares, waterfronts, greenways, and playgrounds, among others. As fundamental components of urban infrastructure, they contribute not only to the physical organization of cities but also to their social, environmental, cultural, and economic vitality. Well-designed and well-managed POSs provide opportunities for recreation, physical activity, social interaction, cultural expression, environmental restoration, and civic engagement, thereby improving urban livability and quality of life [1–6]. They also perform important ecological functions by mitigating urban heat, improving air quality, supporting biodiversity, managing stormwater, and strengthening resilience to climate change [2,6]. In addition, POSs can enhance property values, attract investment and tourism, and support the competitiveness and identity of cities [4,5]. For these reasons, the quality of POSs has become a central concern for architects, planners, designers, policy makers, and local authorities seeking healthier, more sustainable, and more inclusive urban environments.

Because POSs perform multiple and interrelated functions, their quality cannot be understood only through physical appearance or design aesthetics. Public open space quality is a multidimensional construct that reflects the extent to which a space satisfies environmental, functional, social, experiential, and managerial objectives while responding to user needs and expectations [1–6]. This multidimensional conceptualization draws on frameworks that explicitly decompose public space quality into discrete, empirically groundable dimensions rather than treating it as a single undifferentiated attribute [2,4]. Commonly cited quality attributes include accessibility, safety, comfort, maintenance, functionality, inclusiveness, environmental performance, aesthetics, walkability, sociability, vitality, ecological quality, and user satisfaction [1,2,4–6]. However, these attributes are not defined or prioritized consistently across disciplines. Urban planning, urban design, architecture, landscape architecture, environmental psychology, placemaking, public health, and sustainability studies each emphasize different dimensions of POS quality. This disciplinary diversity enriches the field, but it also means that no single, universally accepted conceptualization of POS quality currently exists.

The multidimensional nature of POS quality makes systematic assessment essential. Without transparent methods for evaluating public open spaces, architects, planners, and policy makers cannot reliably diagnose deficiencies, compare alternatives, monitor improvements, or support evidence-based decisions. Consequently, recent decades have witnessed a proliferation of indicator-based POS assessment tools developed for planning, design, management, policy, and research purposes [1–7]. These tools translate complex qualities, such as accessibility, safety, thermal comfort, inclusiveness, ecological performance, and social vitality, into measurable indicators that can support evaluation and comparison. Similar indicator-based assessment systems have emerged across the built environment, including neighborhood sustainability, healthy cities, healthy workplaces, green buildings, inclusive environments, urban resilience, and place-based health [8–10]. Collectively, such tools now function as important decision-support instruments within architecture, urban design, planning, and governance.

Despite this growth, POS assessment remains conceptually fragmented and methodologically heterogeneous. Existing tools have been developed across multiple disciplinary and institutional settings and therefore differ in their theoretical assumptions, construct definitions, measurement structures, and intended applications [1–7,11,12]. While such diversity is partly necessary because public spaces differ by type, culture, climate, and governance context, it creates serious comparability challenges. Two tools applied to the same space may produce different interpretations because they define quality differently,

prioritize different dimensions, rely on different evidence sources, or address different user groups. These differences complicate tool selection, weaken the interpretability of assessment results, and limit the cumulative development of knowledge across studies.

Much existing criticism of POS assessment tools has focused on indicator selection and operationalization. Tools differ substantially in the number, type, and organization of indicators used to represent POS quality. Similar concepts are often expressed through different labels, while identical indicator names may refer to different constructs or measurement procedures. Variation also exists in the extent to which tools report operational definitions, scoring protocols, aggregation procedures, and weighting schemes, reducing transparency and limiting replication [11,12]. This pattern of incomplete and inconsistent indicator coverage is not unique to POS assessment: A structured coverage analysis of urban quality-of-life tools has similarly found a lack of consensus on the type, number, and selection method of indicators across instruments [13]. As a result, assessment outputs may be difficult to interpret or compare, especially when tools are transferred across geographical, disciplinary, or institutional contexts. These inconsistencies have direct implications for content validity, construct interpretation, and the credibility of assessment outcomes for practitioners and decision makers.

Comparative studies have made important contributions by examining the wide range of POS quality dimensions, the comprehensiveness of indicator sets, thematic overlap, indicator frequency, and content validity [4–7]. These studies demonstrate that the field lacks consensus about which dimensions and indicators should define a comprehensive POS assessment framework. However, most comparative work implicitly treats assessment tools as collections of indicators or checklists. This indicator-centered view is useful but incomplete. POS assessment tools also incorporate measurement procedures, hierarchical structures, operational definitions of indicators, intended users, application contexts, and dissemination. These components shape how quality is represented, interpreted, and used. Therefore, tools with similar indicators may still produce non-equivalent assessments if they differ in measurement logic, construct definition, intended use, or institutional authority.

Addressing this limitation requires an analytical perspective that examines POS assessment tools as integrated socio-technical systems rather than isolated indicator lists. Semiotic theory provides such a perspective by conceptualizing assessment tools as layered sign systems that generate meaning through relationships among measurement, structure, interpretation, use, and institutional context. From this standpoint, indicators are only one level of representation. The meaning and practical authority of an assessment output also depend on how indicators are operationalized, how they are organized, what construct they are claimed to represent, who is expected to use the tool, and how the tool gains legitimacy in professional or governance settings. A semiotic perspective therefore offers a broader way to understand divergence while preserving the contextual flexibility required in POS assessment.

Accordingly, this study examines divergence among POS assessment tools through a five-level semiotic framework that conceptualizes assessment instruments as socio-technical sign systems. This study has three objectives: (1) to examine how POS assessment tools differ across empiric, syntactic, semantic, pragmatic, and social semiotic levels; (2) to identify the structural, conceptual, operational, and institutional characteristics associated with divergence among existing tools; and (3) to propose a semiotic compliance framework that specifies minimum documentation requirements for improving transparency, internal coherence, interpretability, and comparability in future POS assessment tools. By shifting attention from indicator diversity alone to the broader architecture of assessment systems, this study contributes a more integrated framework for understanding divergence and

offers practical guidance for researchers, practitioners, and policy makers involved in developing, selecting, and applying POS assessment tools.

2. Methodology

This study adopts a multi-method research design combining a systematic literature review, a structured content analysis, and a semiotic framework synthesis to examine existing public open space (POS) assessment tools. The research process is organized into four sequential phases, as illustrated in Figure 1: (1) tool identification through systematic literature review, (2) identification and extraction of tool attributes through structured content analysis, (3) semiotic framework analysis and synthesis, and (4) statistical analysis.

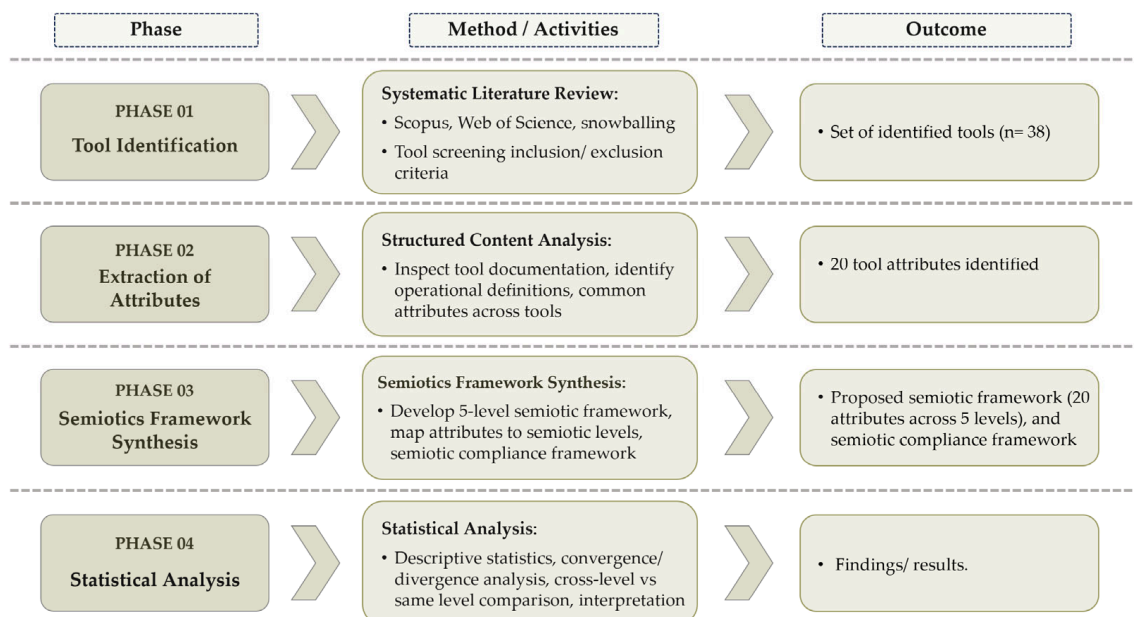


Figure 1. Analytical workflow of this study.

Tool Identification. A systematic literature review was conducted to identify existing tools dedicated to assessing and measuring the quality of POSs. A systematic literature review is a structured, transparent, and rigorous approach to identifying, critically evaluating, and synthesizing the literature addressing a specific research question [14]. The search was conducted on two well-known databases, Scopus and Web of Science. The following two search strings were used to identify all literature related to the topic: The following two strings were used: (“public open space*” OR “urban open space” OR “campus open space” OR “residential open space” OR “placemaking”), and (“assessment tool*” OR “rating tool*” OR “assessment system*” OR “rating system*” OR “assessment framework*” OR “assessment instrument*” OR “evaluation framework” OR “indicator*”). Only publications written in English were included, and no publication date restrictions were applied. The searches initially retrieved 257 records from Scopus and 246 from Web of Science. The retrieved records included journal articles, conference papers, and book chapters. The records were subsequently screened using predefined selection criteria, resulting in 34, as shown in Figure 2. A snowballing procedure was additionally applied to trace relevant tools cited within identified publications, ensuring comprehensive coverage beyond the initial database search. This search resulted in another 4 tools.

An assessment tool was included when it meets the following two conditions: (a) it is based on a well-defined measurement structure comprising categories, subcategories, and indicators; (b) it provides sufficient information and documentation about the tool, its measuring structure, and various relevant details to allow in-depth analysis of the tool.

Tools lacking these two conditions were excluded. The application of these selection criteria resulted in a set of thirty-eight POS assessment tools.

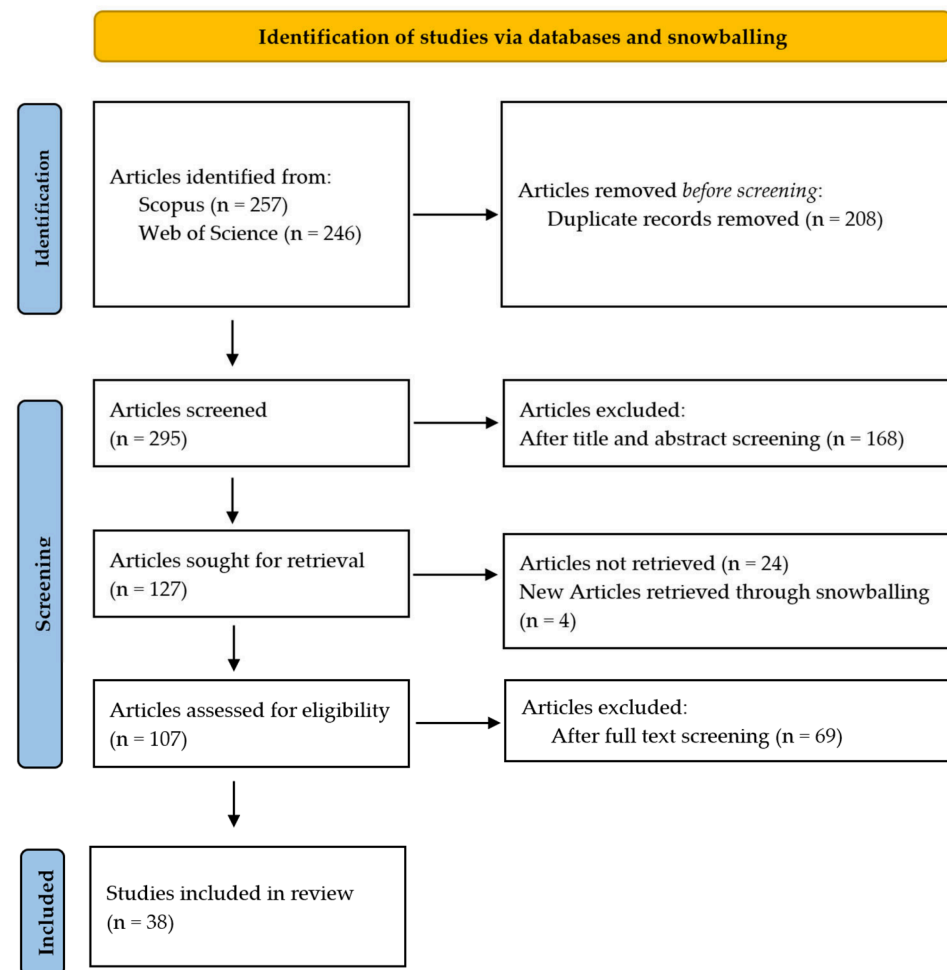


Figure 2. PRISMA flow diagram of the systematic literature review process.

Identifying Attributes of Tools. Structured content analysis was used to systematically identify and extract relevant information about the attributes of each of the identified thirty-eight tools. Content analysis is a well-established method for the systematic examination and interpretation of textual documents to reveal recurring patterns, themes, and underlying characteristics [15]. For each tool, published reports and related documentation were reviewed in full using a standardized coding template. Both deductive and inductive content analyses were employed. Deductive content analysis was guided by a predefined set of analytical codes derived from the literature on the quality of POS. On the other hand, inductive content analysis was applied to identify emerging themes and categories not captured in the predefined coding scheme. All attributes and their possible values, whether identified a priori or generated during coding, were operationally defined before coding began and were applied consistently across all tools. Where documentation was ambiguous, conservative classification was applied based only on verifiable evidence.

To check the reliability of the process, a sample of ten tools was independently coded by two researchers; disagreements were resolved through structured discussion to refine the attributes and their possible values, with an initial inter-rater agreement exceeding 88% prior to reconciliation. This phase resulted in identifying twenty tool attributes. In the subsequent phase, these attributes were mapped to one level of the semiotic framework.

Semiotic Framework Synthesis. A theory-driven analytical framework was developed and used as the interpretive and analytical foundation of this study. Drawing on semiotic theory [16,17], a five-level semiotic framework was developed and applied to examine POS assessment tools as layered socio-technical sign systems. Based on existing semiotic theory as introduced in Section 2, the proposed framework consists of empiric, syntactic, semantic, pragmatic, and social levels, providing an organizing logic that links measurement architecture to meaning, use, and institutional context. This framework enabled us to examine various assessment tools as layered socio-technical sign systems, and thus, all tool attributes and characteristics can be mapped and linked to one of the five semiotic levels of the semiotic framework. The twenty attributes identified in the content analysis phase were mapped, interpreted, and analyzed across the five theory-based levels, making both internal coherence and cross-level tension visible within and across tools.

Statistical Analysis. The final phase of this study involved statistical analyses of the tool attributes extracted and classified in the content analysis phase, and they are synthesized and mapped to the five semiotic levels developed in Phase 3. Descriptive statistics were computed to summarize attribute distributions across the thirty-eight tools, providing frequency distributions for each attribute across the five semiotic levels. Convergence and divergence analysis was conducted to identify where tools converge and where they diverge, both within individual semiotic levels and across levels. Cross-level comparisons were prioritized, as the central analytical claim of this study is that comparability problems in POS assessment arise primarily from cross-level structural misalignment rather than from indicator diversity alone.

3. Semiotic-Based Analysis Framework

To move beyond a reductionist treatment of POS tools as collections of indicators, this study uses semiotic theory to examine these tools as socio-technical sign systems. In practice, POS tools do more than combine indicators into composite scores. They represent complex spatial and social phenomena, attach meaning to constructs such as “quality,” “comfort,” or “publicness,” and communicate those meanings to specific audiences. They are technical because they rely on formal measurement procedures, data infrastructures, and aggregation structures. They are social because their outputs are interpreted, contested, legitimized, and acted upon within professional and political environments.

From a semiotic standpoint, POS assessment tools can be understood as structured information systems that capture, process, and communicate data through measurable indicators [17–19]. Inputs, including statistical records, survey responses, and administrative or captured datasets, are translated into outputs such as POS quality scores or rankings through formalized procedures [11]. These outputs depend on representational elements, including numerical values, labels, categories, and graphical symbols, that stand in for complex spatial realities [19,20]. In semiotic terms, such elements function as signs: observable forms that represent underlying concepts [16,21]. A sign consists of a signifier—the observable form, such as a numerical score—and a signified, the concept to which the signifier refers, such as accessibility, comfort, or quality. The link between the two is not natural but is established through operational definitions, rules, disciplinary conventions, and institutional norms [16,17,20–22]. When a public open space is assessed and assigned a quality score, that value derives meaning from the structured system that defines how it was measured, combined, and interpreted.

Semiotic theory also recognizes that non-linguistic systems, including artifacts, images, and spatial configurations, function as sign systems governed by cultural and institutional codes [23–25]. The analysis of how such systems generate meaning is essential for understanding their operation in the world. In classical theory, any sign operates across

three levels, namely the syntactic, semantic, and pragmatic, which deal, respectively, with structure, meaning, and use [21,26]. Stamper [16] extends this with a social level that addresses the broader consequences of sign systems in their institutional and cultural context. Building on this tradition, the present framework adds an empiric level that addresses the foundational measurement practices through which raw data are transformed into indicator values before structural aggregation. The resulting five-level structure provides an integrative logic that links measurement architecture to meaning, use, and institutional context.

Empiric Level. The empiric level concerns how indicators are specified, operationalized, measured, and calculated before aggregation. It addresses the translation of observed phenomena into quantifiable values through defined measurement procedures and focuses on the technical specification and consistency of measurement. At this level, the system establishes its measurement foundation, ensuring that data are captured and transformed consistently before they are structured, interpreted, or applied [16,27,28].

Syntactic Level. The syntactic level deals with the formal structure of signs and the rules that govern their arrangement. It addresses how symbols relate to one another, independently of their meaning or use, in much the same way that grammar concerns the well-formedness of a sentence rather than what it expresses. In assessment systems, the syntactic level concerns the formal architecture through which indicators are arranged and combined, including the depth of the hierarchy, the number and breadth of categories and subcategories, and the rules governing aggregation into higher-order outputs. A central distinction at this level is between fully operational tools, which prescribe structured workflows and aggregation procedures, and adaptable frameworks, which define domain logics while permitting flexibility in local implementation [16,17].

Semantic Level. The semantic level concerns meaning. It examines how signs relate to the objects or ideas they represent, the signified, and how these relations are made consistent enough to support reliable communication [21]. Meaning is shaped by disciplinary, theoretical, cultural, and professional contexts, linking signs to their real-world referents [16]. In POS assessment, the semantic level examines how constructs such as “POS quality,” “accessibility,” “comfort,” or “social life” are defined and stabilized within a tool, how their conceptual boundaries and relationships are configured, and how disciplinary orientation and indicator provenance shape their meaning. Indicators derived from the academic literature, design guidelines, standards, or sustainability rating systems may embody different assumptions and priorities regarding what constitutes POS quality. Semantic aspects are particularly consequential because tools that appear to address the same domain may define it in fundamentally different ways and therefore yield non-equivalent measurement structures even when their indicator labels overlap.

Pragmatic Level. The pragmatic level addresses how signs are actually used and how they influence behavior, corresponding to what Morris originally defined as the relation of signs to their interpreters [21]. It concerns the relationship between signs and the people who interpret and act on them, including their intentions, decisions, and contexts of action; in organizational information systems, this relationship is governed by the social norms through which signs acquire the authority to guide action [29]. Context is central at this level, because the same sign can be used differently depending on who uses it and for what purpose [17]. In the POS assessment, pragmatic analysis examines the intended users of a tool, the spatial scale at which it is designed to operate, and the type or typological scope of public open space to which it is intended to be applied. Different application contexts, including civic, residential, and campus open spaces, may involve different users, spatial scales, typological scopes, and decision purposes. These differences make the intended context of application central to the pragmatic interpretation of an assessment tool.

Social Level. The social level considers the broader institutional and cultural context in which signs operate. It focuses on the shared values, norms, conventions, and governance structures that shape how signs are interpreted, trusted, and acted upon within society. While assessment systems are designed with intended purposes, their influence ultimately depends on the institutional settings and cultural contexts in which they are developed, adopted, and applied [16,29]. In POS assessments, the social level examines the institutional origins, development process, geographic and cultural context, operational embedding, and legitimacy pathways of assessment tools. These pathways include scholarly review, stakeholder participation, institutional adoption, and the transparent reporting of reliability and validity evidence, all of which contribute to users' confidence in an instrument. Research-based tools typically establish legitimacy through scholarly argumentation and peer review, whereas operational tools often derive authority from institutional adoption and regulatory recognition. The extent to which these legitimacy mechanisms are documented, however, varies considerably among tools and should not be inferred solely from their research-based or operational status. Although validity concerns the relationship between an assessment and the construct it measures, the present framework treats reported reliability and validity evidence as social attributes because the analysis focuses on how such evidence supports the external credibility, trustworthiness, and legitimate use of assessment tools. Figure 3 summarizes the five-level semiotic framework described above.

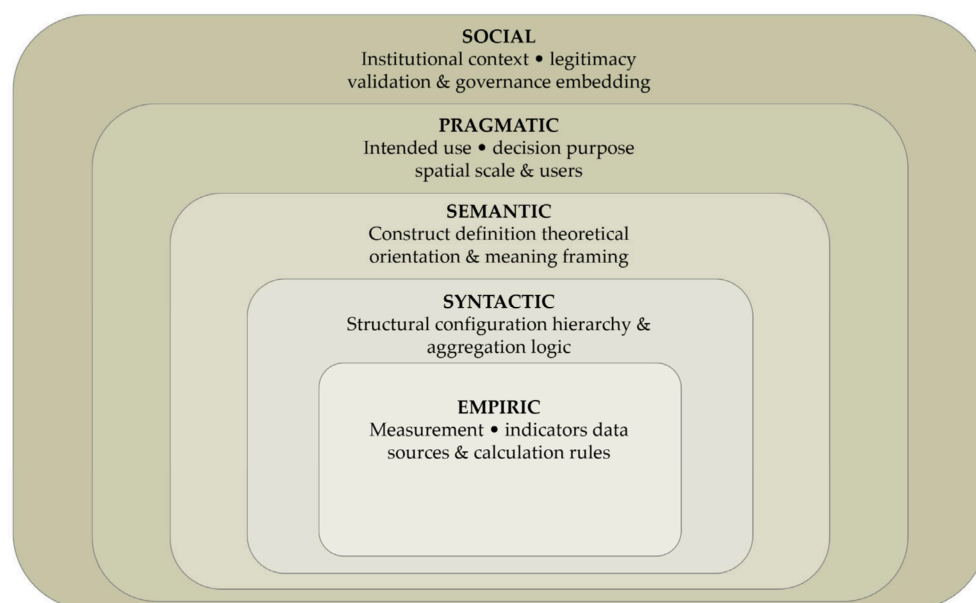


Figure 3. Five-level semiotic framework for examining POS assessment tools.

4. Twenty Semiotic Attributes of POS Assessment Tools

Applying the methodological approach described above, the analysis of thirty-eight POS assessment tools identified twenty recurring attributes that characterize how these tools operate across the five semiotic levels. These attributes were derived from structured content analysis of official documentation and methodological reports. They are empirical features of the tools themselves and not theoretical assumptions imposed on the data.

The twenty attributes capture observable features of tool design, including how indicators are defined and operationalized, how assessment structures are organized, how constructs are conceptualized and bounded, how tools are positioned within decision processes, and how institutional authority is established. Table 1 summarizes the twenty attributes by semiotic level.

Table 1. Semiotic analysis framework: twenty attributes of POS assessment tools mapped to semiotic levels.

Semiotic Level	No. & Attribute	What the Attribute Captures
Empiric	1. Type of indicators	Whether data are objective, subjective, or a combination of both and the proportion of each type.
	2. Availability of operational definition (for objective indicators)	Does the tool provide an operational definition for indicators?
	3. Question stems reported (for subjective indicators)	Whether question stems of the assessment tool surveys are properly worded, standardized, and reported for subjective indicators.
	4. Indicator clarity & specificity	The degree to which each indicator is formally defined in measurable terms, specifying what is measured and how its value is derived.
	5. Indicator intent specified	Whether the purpose or objective of each indicator is explicitly stated in the tool documentation.
Syntactic	6. Level of detail and sophistication	Whether the system is a complete, detailed assessment instrument (i.e., a tool), a partial assessment instrument (i.e., a framework), or a set of assessment indicators.
	7. Depth of hierarchical structure	The number of measurement levels in the tool's hierarchical structure (depth of coverage).
	8. Number of categories	The number of categories in the measurement structure (i.e., breadth of coverage at the category level).
	9. Number of sub-categories	The number of categories in the measurement structure (i.e., breadth of coverage at the sub-category level).
Semantic	10. Field of study	The field of study and the theoretical orientation in which the tool and its POS construct are grounded.
	11. Construct measured	The core concept or construct that the tool aims to assess or measure.
	12. Source of indicators	The source from which the indicators were extracted and adapted, such as academic literature, policy, design guidelines, or existing assessment tools.
Pragmatic	13. Intended users	The target or expected users of the tool (ordinary people, professionals, or experts).
	14. Spatial scale of application	The urban or spatial scale at which the tool is intended to be used: for instance, national level, city level, or part of a city, such as a neighborhood, district, or university campus.
	15. Type of POS	The type of POS settings for which the tool is designed. For instance, some tools are designed to assess parks (dominated by softscape), plazas, streets, or a combination of them.
Social	16. Operational vs. research-based	Whether the tool is used in real-world practice or developed in academic research and serves primarily academic purposes; related to the legitimacy pathway through which the tool gains authority.
	17. Reliability testing evidence	Whether reliability testing or replication guidance is reported.
	18. Validity testing evidence	Whether validity testing/ argumentation is reported.
	19. Development process	Whether the tool was developed by individual researchers, collaborative academic teams, or institutional bodies.
	20. Geographic and cultural context	The geographic region and cultural setting in which the tool was developed.

This mapping of attributes across the five semiotic levels provides the analytical foundation for the results that follow. The aim is not to evaluate tools as superior or inferior

but to examine how tool attributes located at different levels generate a wide range of misalignment and patterned divergence among tools. By situating empirically identified features within a layered semiotic structure, the analysis makes cross-level misalignment visible within POS assessment tools and supports internal coherence. Recognizing such misalignment does not imply that divergence among tools is inherently problematic. Rather, it allows legitimate variation rooted in disciplinary or governance context to be distinguished from irrational divergence arising from implicit or inconsistent cross-level commitments.

5. Results and Discussion

5.1. Examined Tools

A total of thirty-eight public open space (POS) assessment tools were identified in existing literature and constitute the analytical dataset for this study (Table 2). Collectively, these tools represent a broad cross-section of contemporary approaches to POS quality assessment, reflecting considerable diversity in their disciplinary origins, geographic and cultural contexts, assessment purposes, intended users, application settings, and methodological characteristics. They were developed across multiple countries and academic traditions, addressing different types of public open spaces, and operate at various spatial scales ranging from individual sites to broader urban contexts. The tools also differ in their assessment approaches, incorporating objective indicators, subjective evaluations, or combinations of both, and vary in their structural complexity, ranging from conceptual frameworks to fully operational assessment instruments. This breadth demonstrates the richness and heterogeneity of the current POS assessment landscape and provides a robust empirical basis for examining how divergence emerges across multiple dimensions of assessment design. The following sections analyze these tools systematically through the proposed five-level semiotic framework.

Table 2. The 38 examined POS assessment tools.

ID	No of Ind.	Subjective Indicators	Objective Indicators	No Categories	No of Subcategories	Source
T01	8	4	4	3	8	Strydom & Puren [30]
T02	13	6	7	3	13	Switalski et al. [31]
T03	9	2	7	6	9	John et al. [4]
T04	11	3	8	3	3	Santos Nouri & Costa [32]
T05	40	12	28	8	40	Tawfik Mohamed et al. [6]
T06	15	10	5	4	15	Huang et al. [2]
T07	31	31	0	6	31	Islam et al. [3]
T08	32	16	16	3	15	Skarzauskiene & Maciulienė [33]
T09	19	0	19	7	19	Addas [34]
T10	32	11	21	4	26	PPS [35]
T11	60	18	28	2	11	Lorenzo et al. [36]
T12	49	12	23	8	49	Praliya & Garg [5]
T13	90	18	28	11	90	Knobel et al. [37]
T14	49	49	0	5	15	Mela et al. [38]
T15	13	13	0	3	13	Uysal Bilge [39]
T16	16	16	0	5	16	Asfour & Hossain [40]

Table 2. *Cont.*

ID	No of Ind.	Subjective Indicators	Objective Indicators	No Categories	No of Subcategories	Source
T17	30	15	15	5	30	Imam et al. [41]
T18	14	14	0	1	14	Addas & Maghrabi [1]
T19	18	9	9	3	8	Wijaya & Christianiti [42]
T20	16	16	0	2	16	Aydin & Ter [43]
T21	24	24	0	3	24	Salama [44]
T22	50	50	0	4	50	Khassawneh & Khasawneh [45]
T23	15	15	0	2	11	Faragallah [46]
T24	35	35	0	3	35	Abbasi et al. [47]
T25	12	12	0	2	12	Soliman et al. [48]
T26	5	24	0	2	5	Zohra [49]
T27	9	9	0	1	9	Ernawati et al. [50]
T28	10	10	0	4	64	Haddad [51]
T29	12	12	0	6	31	Abu-Ghazalah & Al-Goussous [52]
T30	13	13	0	5	13	Addas & Maghrabi [53]
T31	18	18	0	4	18	Jamhawi et al. [54]
T32	17	17	0	4	19	Thilakaratne [55]
T33	24	24	0	10	24	Kozlova & Kozlov [56]
T34	10	10	0	6	10	Rad & Ngah [57]
T35	45	45	0	5	13	Ahirrao & Khan [58]
T36	38	0	38	2	38	Diaz-Sarachaga et al. [59]
T37	8	0	8	5	8	Salati et al. [60]
T38	29	0	29	7	29	Hajr et al. [61]

5.2. Empiric-Level Tool Attributes

At the empiric level, the analysis examines how indicators are defined, measured, and calibrated before structural aggregation, focusing on the procedural specification of indicators, the data on which they rely, and the consistency of their treatment within each tool. Five attributes are assessed using Table 3 as the reference.

Type of Indicators. Subjective-only tools form the largest group at 55.3% (21 of 38), followed by mixed-type tools at 34.2% (13 of 38) and objective-only tools at 10.5% (4 of 38). This composition is not incidental: The seven campus tools identified in Section 5.1 are all subjective-only instruments built around occupant perception surveys, and their inclusion in the identified tool set is the principal reason subjective measurement predominates so strongly. This has a direct methodological implication for cross-tool comparison: A subjective-only perception survey and an objective-only physical audit are not measuring the same evidentiary object even when both are labelled POS quality instruments, so any comparison between them must first establish whether they are meant to capture experience or physical condition, rather than treating the two as interchangeable.

Table 3. Empiric-level attributes of the POS assessment tools.

Attribute	Category	No. of Tools	(%)
1. Type of Indicators	Subjective only	21	55.3%
	Objective only	4	10.5%
	Mixed (objective and subjective)	13	34.2%
2. Availability of Operational Definition	Yes	16	42.1%
	Partial	8	21.1%
	No	14	36.8%
3. Question Stems Reported	Yes	21	55.3%
	No	16	42.1%
	Partial	1	2.6%
4. Indicator Clarity & Specificity	Not applicable (0)	14	36.8%
	Not clear (1)	8	21.1%
	Somehow clear (2)	13	34.2%
	Clear (3)	3	7.9%
5. Indicator Intent Specified	Yes	37	97.4%
	No	1	2.6%

Availability of Operational Definition. Sixteen tools (42.1%) fully specify what is measured and how each indicator value is derived, a further eight (21.1%) provide partial definitions, and fourteen (36.8%) provide none, offering only conceptual intent statements without measurement specification. In aggregate, therefore, close to three in five tools (57.9%) fall short of full procedural specification. T36, T37, and T38 are consistently among the most fully specified instruments in the identified set; the practical implication is that procedural specification in POS assessment appears to vary independently of disciplinary convention, reflecting tool-specific documentation practice rather than a single dominant driver.

Question Stems Reported. Twenty-one tools (55.3%) fully report the question stems used to elicit subjective data, while the remaining sixteen (42.1%) either omit them entirely or do not report them clearly enough to be verified, and one tool (2.6%) reports them only partially. Because subjective indicators depend entirely on how a question is worded to elicit a comparable response, the fact that more than four in ten tools do not verifiably report their question stems introduces a reproducibility gap that is easy to overlook: two tools can report an identical indicator label, such as perceived safety, while asking respondents entirely different questions, producing outputs that are nominally but not operationally comparable.

Indicator Clarity and Specificity. Using the four-level clarity scale (0—not applicable; 1—not clear; 2—somehow clear; 3—clear), fourteen tools (36.8%) are coded not applicable because they lack an operational definition altogether, eight (21.1%) are not clear, thirteen (34.2%) are somehow clear, and three (7.9%) achieve full clarity. Read alongside the operational-definition finding above, as this shows that the specification problem in POS assessment is concentrated at the front end of the measurement chain: once a tool commits to defining its indicators at all, it tends to do so with at least partial clarity, but more than a third of tools never reach that starting point.

Indicator Intent Specified. Thirty-seven tools (97.4%) explicitly state the purpose or objective of each indicator; only one tool (T34) does not. Near-universal intent specification

provides a shared baseline across the tool set, but it is a necessary rather than a sufficient condition for measurement transparency: stating why an indicator exists does not substitute for specifying how it is measured, which is precisely where Table 3 shows the tool set to be weakest.

Taken together, the empiric-level findings indicate that POS assessment tools differ far less in their stated commitment to measuring quality than in the procedural rigor with which that commitment is executed. Formal intent is nearly universal, but operational follow-through, definitions, question stems, and indicator clarity are inconsistent and, for well over half the tool set, incomplete. This empiric-level fragmentation establishes a baseline instability that propagates into every higher semiotic level examined below: a tool cannot be structurally, semantically, or pragmatically coherent if its indicators are not first measured consistently.

5.3. Syntactic-Level Attributes

At the syntactic level, attention turns from measurement to structural architecture: how empirically defined indicators are organized, grouped, and transformed into composite outputs. Table 4 reports the distribution of the two syntactic-level attributes across the thirty-eight tools.

Table 4. Syntactic-level attributes of the POS assessment tools.

Attribute	Category	No. of Tools	(%)
6. Level of Detail and Sophistication	Assessment Framework	21	55.3%
	Assessment Tool	17	44.7%
7. Depth of Hierarchical Structure	2 levels (L1–L2)	27	71.1%
	3 levels (L1–L3)	11	28.9%
8. Number of Categories	Minimal (1–4)	22	57.9%
	Moderate (5–8)	14	36.8%
	High (9–12)	2	5.3%
9. Number of Sub-Categories	Minimal (1–20)	24	63.1%
	Moderate (21–50)	12	31.6%
	High (51–100)	2	5.3%

Level of Detail and Sophistication. Twenty-one tools (55.3%) are coded as assessment frameworks, defining domain logics and indicator categories while leaving room for contextual adaptation, and seventeen (44.7%) are coded as assessment tools, presenting a more fully elaborated and procedurally closed measurement instrument. This near-even split indicates that POS assessment has not converged on a single preferred level of structural elaboration: roughly as many developers leave interpretive latitude for local adaptation as developers specify a complete, ready-to-apply instrument.

Depth of hierarchical structure. Twenty-seven tools (71.1%) organize indicators across two hierarchical levels (L1–L2), supporting structured aggregation from indicators to thematic domains, while the remaining eleven (28.9%) extend to a third level (L1–L3), enabling greater measurement granularity. No tool in the identified set relies on a single, flat indicator list. The practical implication is that hierarchical aggregation, and not merely indicator selection, is a near-universal design choice in POS assessment, meaning that comparability problems between tools are as likely to arise from how indicators are grouped and weighted across levels as from which indicators are chosen in the first place.

Number of Categories. Most tools organize their indicators into a compact category structure: 57.9% (22 of 38) use a minimal set of one to four categories, 36.8% (14 of 38) adopt a moderate structure of five to eight categories, and only 5.3% (2 of 38) extend to a high category count of nine to twelve. This concentration at the minimal-to-moderate end indicates that a compact thematic structure is the prevailing design choice and that the small subset of tools with substantially more elaborated category schemes constitutes a structural outlier rather than a common pattern.

Number of Sub-Categories. A similar pattern holds at the sub-category level: 63.1% of tools (24 of 38) specify a minimal range of one to twenty sub-categories, 31.6% (12 of 38) adopt a moderate range of twenty-one to fifty, and only 5.3% (2 of 38) extend to a high range of fifty-one to one hundred. Read together with the category-level findings, this indicates that breadth of coverage, like depth, is concentrated at the lower end of the observed range for most tools, with a small number of substantially more granular instruments setting the upper bound rather than representing a typical design choice.

Overall, the syntactic-level findings show that POS tools converge much more strongly on the principle of hierarchical aggregation than on the degree of procedural closure they adopt. This distinction matters for comparability: Two three-level tools are not necessarily more alike than a two-level and a three-level tool, because the number of levels governs measurement granularity, whereas the framework versus tool distinction governs how much interpretive discretion is left to the assessor. Both axes of variation operate independently of indicator content and constitute a source of divergence that persists even when tools nominally measure the same constructs. The number of categories and sub-categories adds a third, largely independent axis of syntactic variation: most tools converge on a compact category and sub-category structure so that the small minority of more granular instruments diverge from the rest of the tool set on breadth of coverage, even where they share the same depth and procedural-closure profile.

5.4. Semantic-Level Attributes

At the semantic level, the analysis examines how POS assessment tools construct and communicate meaning: how POS quality and its sub-dimensions are conceptually defined and tied to real-world phenomena. Table 5 reports the distribution of semantic-level attributes across the thirty-eight tools.

Table 5. Semantic-level attributes of the POS assessment tools.

Attribute	Category	No. of Tools	(%)
10. Field of Study	Urban design	14	36.8%
	Urban planning	13	34.2%
	Architecture	4	10.5%
	Landscape architecture	1	2.6%
	Architecture & urban design	6	15.8%
11. Construct Measured	Public open space quality	33	86.8%
	Open space quality (university campus)	5	13.2%
12. Source of Indicators	Literature (on POS assessment tools, models, indicators, & metrics)	30	78.9%
	Design guidelines & standards	5	13.2%
	Sustainable neighborhood rating systems	3	7.9%

Field of Study and Theoretical Orientation. As established in Section 5.1, the tool set concentrates in four related design and planning traditions, led by urban design (36.8%, 14 of 38) and urban planning (34.2%, 13 of 38), with architecture and urban design (15.8%, 6 of 38), architecture alone (10.5%, 4 of 38), and landscape architecture (2.6%, 1 of 38) accounting for the remainder. Each field functions as an interpretive frame that determines whether POS quality is approached primarily as a spatial-configurational problem, a governance and land-use problem, or a material and ecological one, with each framing producing distinct measurement priorities even where tools nominally address the same public spaces.

Construct Measured. Semantic divergence is reinforced by the primary construct each tool is designed to assess. Public open space quality is by far the dominant construct, accounting for 86.8% of tools (33 of 38), confirming that the large majority of the identified tools share a common conceptual target even where their indicators, structures, and intended uses diverge. The remaining 13.2% (5 of 38) target open space quality specifically within a university campus context, a construct that retains the general emphasis on spatial quality but is bounded by institutional academic governance rather than civic or municipal governance. This construct distinction is semantically consequential even though it separates only a small share of the tool set: it confirms that the campus sub-cluster identified in Section 5.1 is not merely a pragmatic or geographic outlier but a construct-level distinction in its own right, one that should be made explicit whenever campus-oriented and civic POS tools are compared.

Source of Indicators. Indicator development is overwhelmingly grounded in the existing literature: 78.9% of tools (30 of 38) were developed based on the existing literature on indicators, with authors adopting and adapting metrics, measures, and indicators already reported in the literature rather than generating them independently. Design guidelines and standards account for a further 13.2% (5 of 38), and sustainable neighborhood rating systems, the indicator source associated with T36, T37, and T38, account for the remaining 7.9% (3 of 38). The near-four-fifths concentration on literature-based sourcing carries an important implication for construct validity: indicators derived primarily from the academic literature may not adequately capture the POS quality dimensions most meaningful to local communities and users, a concern reinforced by how little of the tool set draws directly on standards, policy documents, or dedicated primary data collection.

The semantic-level findings indicate that divergence among POS tools is rooted in how meaning is constructed through disciplinary framing, construct definition, and indicator provenance rather than in measurement technique alone. Tools that appear comparable because they share indicator labels or measurement structures may nonetheless be semantically non-equivalent, which helps explain why nominally similar tools can produce different interpretations of what POS quality means in practice.

5.5. Pragmatic-Level Attributes

At the pragmatic level, the analysis addresses how POS assessment tools are intended to function in practice: the actors they serve and the spatial and typological contexts in which they are designed to operate. Pragmatic differences do not concern what a tool's indicators measure, which is addressed at the semantic level, but how its outputs are interpreted, by whom, and within what physical and typological scope, making this level central to whether two tools' results can be meaningfully compared. Table 6 reports the distribution of pragmatic-level attributes.

Table 6. Pragmatic-level attributes of the POS assessment tools.

Attribute	Category	No. of Tools	(%)
13. Intended Users	Users	19	50.0%
	Experts	12	31.6%
	Users & experts	7	18.4%
14. Spatial Scale of Application	Neighborhood/urban areas	29	76.3%
	City-wide	7	18.4%
	National	2	5.3%
15. Type of POS	Parks	6	15.8%
	Parks & plazas	17	44.7%
	Parks, plaza & street	15	39.5%

Intended Users. Users form the largest group, at 50.0% (19 of 38), reflecting the substantial weight of campus-based and residential perception-survey tools organized around direct questionnaire administration rather than expert field auditing. Experts account for 31.6% (12 of 38), representing tools that require professional technical knowledge for indicator scoring and interpretation, while a further 18.4% (7 of 38) are designed for joint use by both users and experts, aiming to triangulate physical conditions with subjective experience. Every tool in the identified set specifies an intended user group; none is designed for unmediated use by the general public in a participatory governance setting, suggesting that POS quality assessment, whether expert-led or user-facing, remains a facilitated activity rather than an open community tool. This division carries a practical consequence for comparability: an expert-oriented tool and a user-facing tool addressing the same public space can reach different conclusions about its quality not because the space itself differs but because each tool channels judgment through a different evaluative actor, one applying technical criteria and the other reporting lived experience.

Spatial Scale of Application. The large majority of tools, 76.3% (29 of 38), are designed for application at the neighborhood or broader urban-area scale, confirming that POS assessment is overwhelmingly conceived as a sub-city, localized activity. City-wide tools account for 18.4% (7 of 38), and only 5.3% (2 of 38) are designed for national-level application. Fewer than one in five tools is designed to operate at a scale broader than the neighborhood or urban-area level, which has a direct practical implication for municipalities seeking to benchmark POS quality city-wide, since most available instruments were not designed with that comparative scope in mind.

Type of POS. Tools designed to assess parks and plazas together form the largest group at 44.7% (17 of 38), followed by tools that extend this scope to include streets as well at 39.5% (15 of 38), covering parks, plazas, and streets together. A smaller group of tools, 15.8% (6 of 38), is restricted specifically to parks. This typological range sharpens the pragmatic non-comparability already implied by the spatial-scale findings: a tool restricted to parks and a tool spanning parks, plazas, and streets are not validated over the same physical or experiential territory, even where both are labelled general POS assessment tools, and cross-tool comparison should account for this difference in typological scope before treating their outputs as equivalent.

Together, the pragmatic-level findings indicate that divergence among POS assessment tools is strongly shaped by who a tool is designed for and where it is designed to be applied. These pragmatic variations help explain why tools that appear similar at the indicator level may nonetheless lead to different actions and outcomes in practice, and they reinforce the

case, developed further in Section 5.7, for situating each tool explicitly within its intended user group, spatial scale, and POS typology before using it for cross-context comparison.

5.6. Social-Level Attributes

At the social level, the analysis examines POS assessment tools in terms of who develops them, their geographic and cultural provenance, and their degree of institutional embeddedness. Table 7 reports the distribution of social-level attributes.

Table 7. Social-level attributes of the POS assessment tools.

Attribute	Category	No. of Tools	(%)
16. Operational vs. Research-Based	Research-based tools	38	100%
	Operational, real-life tools	0	0%
17. Reliability Testing Evidence	Fully reported	2	5.3%
	Partially reported	3	7.9%
	Not reported	32	84.2%
	Unclear	1	2.6%
18. Validity Testing Evidence	Fully reported	1	2.6%
	Partially reported	4	10.5%
	Not reported	29	76.4%
	Unclear	4	10.5%
19. Development Process	Authors	25	65.8%
	Authors & experts	8	21.0%
	Authors, experts, & stakeholders	5	13.2%
20. Geographic and Cultural Context	Europe & UK (Spain, Greece, Turkey, UK, Russia)	8	21.1%
	Middle East & North Africa (Jordan, Qatar, Saudi Arabia, Egypt Algeria)	14	36.8%
	Asia (China, Bangladesh, India, Indonesia)	7	18.4%
	Americas (United States)	1	2.6%
	Not specified	8	21.1%

Operational vs. Research-Based Status. Table 7 classifies all thirty-eight tools, including T36, T37, and T38, as research-based (100%, 38 of 38), reflecting the uniform content-analysis protocol described in Section 2, which was applied to official documentation for every tool in the identified set regardless of institutional origin.

Reliability Testing Evidence. Reporting of reliability testing is rare across the identified tools: Only 5.3% (2 of 38) fully report reliability testing or replication guidance, and a further 7.9% (3 of 38) report it partially. The large majority, 84.2% (32 of 38), do not report reliability evidence at all, and reporting is unclear for the remaining 2.6% (1 of 38). This near-absence of reliability documentation means that, for most tools, a user cannot establish whether repeated application under comparable conditions would yield consistent results, which limits confidence in cross-context or longitudinal use of their outputs.

Validity Testing Evidence. Validity evidence is similarly scarce: only 2.6% (1 of 38) of tools fully report validity argumentation or testing, while 10.5% (4 of 38) report it partially. Validity evidence is not reported for 76.4% (29 of 38) of tools, and it is unclear for a further 10.5% (4 of 38). Taken together with the reliability findings, this indicates that psychometric

validation is largely absent from POS assessment tool documentation, so claims about what a tool measures, and how dependably it does so, generally rest on the tool developers' own conceptual justification rather than on independently reported testing evidence.

Development Process. Authors alone develop nearly two-thirds of the identified tools (65.8%, 25 of 38), producing a body of instruments for which their indicator vocabulary is diverse but fragmented, as successive independently authored tools introduce slightly different constructs for closely related phenomena. Tools developed jointly by authors and experts account for a further 21.0% (8 of 38), and a smaller group, 13.2% (5 of 38), is developed through a broader collaborative process involving authors, experts and users together. This concentration of development activity in individual academic authorship, with collaborative pathways together accounting for barely a third of the tool set, offers a structural explanation for the semantic fragmentation documented in Section 5.4: indicator vocabularies proliferate in part because tool development is rarely subjected to the standardizing pressure of broader stakeholder validation.

Geographic and Cultural Context. The identified tools span fifteen or more national contexts, consolidated in Table 7 into five broad regions. The Middle East and North Africa region, spanning Jordan, Qatar, Saudi Arabia, Egypt and Algeria, contributes the largest specified share at 36.8% (14 of 38). Europe and the United Kingdom, including Spain, Greece, Turkey and Russia, account for 21.1% (8 of 38). Eight tools (21.1%) do not specify a geographic context, implicitly positioning themselves as universally applicable. Asia, spanning China, Bangladesh, India, and Indonesia, contributes 18.4% (7 of 38), and the Americas contribute 2.6% (1 of 38) from the United States. This geographic breadth carries a direct social implication, since cultural and institutional contexts shape what counts as credible evidence, which stakeholder voices are prioritized in tool development and how legitimacy is ultimately constructed, whether through policy alignment, academic validation, or community participation.

Overall, the social-level analysis shows that POS assessment tools differ substantially in how they establish legitimacy, credibility, and trust. Divergence arises not only from methodological choices examined at the empiric and syntactic levels but from who develops a tool, where it is developed and validated, and through which institutional pathway it gains authority. These social factors help explain why comparability challenges persist even where tools appear similar in indicator content or structural form. The scarcity of reported reliability and validity evidence compounds these challenges, since institutional legitimacy and psychometric legitimacy are established through different pathways, and a tool can be institutionally credible while remaining empirically untested in the psychometric sense.

5.7. Divergence Across POS Assessment Tools

The level-by-level analysis in Sections 5.2–5.6 demonstrates that divergence among POS assessment tools is far more pervasive than indicator selection alone can explain. The dominant patterns of variation arise from cross-level semiotic misalignment, particularly at the semantic, pragmatic, and social levels, where meaning, intended use, validation of the tool and institutional legitimacy are constituted. To quantify the degree to which the tools diverge/converge across five semiotic levels, Jaccard similarity analysis was conducted. Pairwise tool similarity was calculated using the set-based (binary) Jaccard similarity index, in which each tool was represented as the set of coded categorical values it exhibits across the twenty semiotic attributes, and similarity was defined as the ratio of shared coded attributes to the union of attributes exhibited by either tool. Figure 4a–f present the resulting Jaccard similarity heatmaps. Table 8 on other hand, summarizes the substantive comparability challenges underlying these patterns.

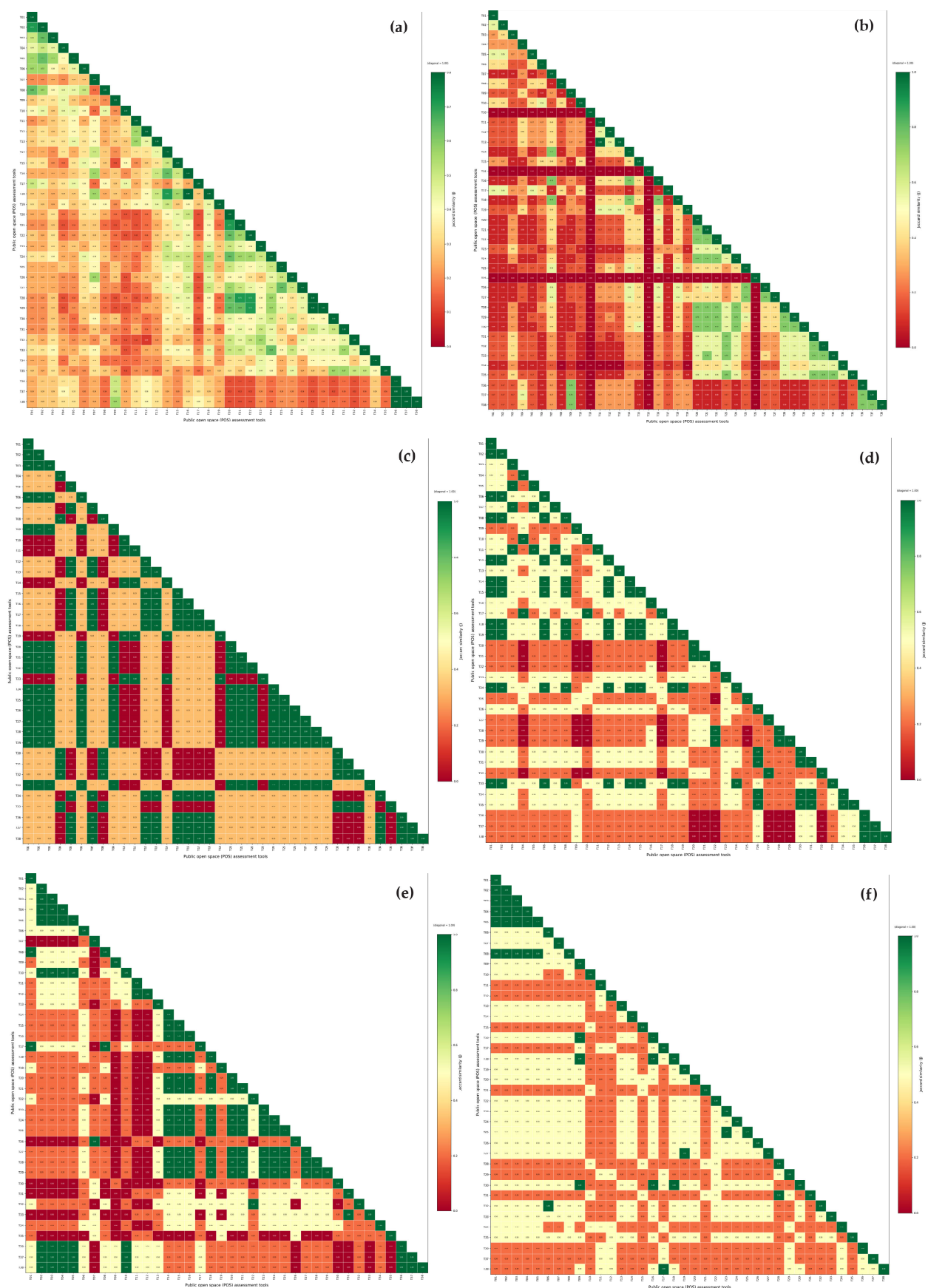


Figure 4. Pairwise Jaccard similarity heatmaps among the 38 identified POS assessment tools at six levels (a) tool level; (b) empiric level; (c) syntactic level; (d) semantic level; (e) pragmatic level; (f) social. The degree of convergence varies markedly across semiotic levels. The empiric level is the

most differentiated (mean $J = 0.24$; median $J = 0.17$): 57.3% of pairs fall below $J = 0.20$, 15.0% share no empiric-level coding in common ($J = 0$), and only one pair (T11 and T13) reaches $J = 1.00$. This pronounced divergence is consistent with the variation documented in evidence type, operational definitions, question-stem reporting, and indicator clarity. Tools may state similar indicator intentions while operationalizing them through substantially different forms of evidence, making apparent content overlap an unreliable basis for comparison.

Table 8. Patterns of divergence across POS assessment tools: semiotic-level comparability challenges.

Semiotic Level	Attribute	Comparability Challenge
Empiric	Type of Indicators	Tools range from entirely objective assessment instruments to entirely subjective perception surveys (55.3% subjective-only); their outputs are not directly comparable without first establishing which evidentiary base each represents.
	Availability of Operational Definition	Over a third of tools (36.8%) provide no operational definition; identical indicator labels may therefore operationalize different underlying constructs across tools.
	Question Stems Reported	Question stems are absent or cannot be verified in 42.1% of tools, preventing consistent replication of subjective data collection.
	Indicator Clarity & Specificity	More than a third of tools (36.8%) cannot be assessed for clarity because no operational definition exists; among all tools, only 7.9% achieve full clarity.
	Indicator Intent Specified	Near-universal (97.4%) but insufficient alone: stating an indicator's purpose does not establish how that indicator is measured.
Syntactic	Level of Detail and Sophistication	Frameworks (55.3%) permit flexible local adaptation while tools (44.7%) impose procedural closure; neither format alone ensures cross-tool comparability.
	Depth of Hierarchical Structure	Two-level (71.1%) and three-level (28.9%) aggregation structures produce outputs that differ in measurement granularity even when built from similar indicators.
	Number of Categories	Category counts range from minimal (57.9%, 1–4 categories) to high (5.3%, 9–12 categories); tools with markedly different category counts may appear structurally comparable while differing substantially in analytical granularity.
	Number of Sub-Categories	Sub-category counts show the same concentration pattern (63.1% minimal, 5.3% high), compounding the category-level granularity mismatch and further limiting like-for-like comparison of coverage breadth.
Semantic	Field of Study	Four disciplinary traditions frame POS quality through distinct theoretical logics, producing non-equivalent construct boundaries for nominally shared indicator domains.
	Construct Measured	The dominant public open space quality construct (86.8%) and the campus-bounded open space quality construct (13.2%) are conceptually distinct, even though both concern spatial quality, because one is framed within civic governance and the other within institutional academic governance.
	Source of Indicators	Literature-sourced indicators (78.9%) reflect an academic epistemological foundation, while design guidelines and standards (13.2%) and sustainable neighborhood rating systems (7.9%) draw on institutional and applied sources; these differing epistemological foundations constrain cross-source comparability.

Table 8. Cont.

Semiotic Level	Attribute	Comparability Challenge
Pragmatic	Intended Users	Tools designed for expert professionals (31.6%) and those designed for general users (50.0%) are built around different evidence bases and decision logics.
	Spatial Scale of Application	Most tools (76.3%) are designed for neighborhood- or urban-area-scale application; city-wide and national tools are comparatively rare, limiting the pool of instruments suited to city-wide benchmarking.
	Type of POS	Tools restricted to parks, tools covering parks and plazas, and tools covering parks, plazas and streets are not validated over the same physical or experiential territory.
Social	Operational vs. Research-Based	All thirty-eight tools are classified as research-based under this study's uniform coding protocol (Section 5.6); this attribute does not differentiate the tool set, so comparability should rely on the other social-level attributes instead.
	Reliability Testing Evidence	Reliability testing or replication guidance is reported, fully or partially, for only 13.2% of tools; for the remaining 86.8%, whether repeated application would yield consistent results cannot be independently verified.
	Validity Testing Evidence	Validity argumentation or testing is reported, fully or partially, for only 13.1% of tools; for the remaining 86.9%, claims about what a tool measures rest largely on developer assertion rather than demonstrated evidence.
	Development Process	Author-led tools (65.8%) generate diverse but fragmented indicator vocabularies; collaboratively developed tools (34.2%) offer more standardization but reduced contextual sensitivity.
	Geographic and Cultural Context	Context-specific climatic, cultural and governance conditions across fifteen or more countries constrain direct indicator transferability; 21.1% of tools do not specify a context at all.

Across the complete semiotic profile, similarity among the 703 unique tool pairs was low to moderate (mean $J = 0.31$; median $J = 0.29$). Only approximately 12% of pairs reached $J \geq 0.50$, whereas approximately 16% fell below $J = 0.20$; no two tools shared an identical overall profile. The highest observed similarity was $J = 0.80$, occurring for five pairs, including the three pairwise combinations among T36, T37, and T38. These three tools form a coherent sub-cluster but remain comparatively distinct from the wider set, with an average similarity of $J = 0.23$ to the other tools. Overall convergence is therefore concentrated in a limited number of pairs and clusters rather than organized around a dominant field-wide assessment model.

Syntactic similarity is the highest of the five levels (mean $J = 0.45$), with 27.9% of pairs sharing identical profiles ($J = 1.00$). This apparent convergence should be interpreted cautiously because the level comprises only two categorical attributes, producing a small number of possible combinations. The result indicates convergence around broad structural forms, particularly hierarchical organization, but does not demonstrate equivalence in scoring, weighting, or aggregation procedures. Structural resemblance at this coarse level can therefore coexist with substantial divergence elsewhere.

Semantic similarity is moderate (mean $J = 0.39$), with 49.1% of pairs reaching $J \geq 0.50$ and only 7.4% sharing no coding. This reflects concentration around the dominant POS-quality construct and literature-based indicator sourcing. Nevertheless, tools sharing a disciplinary field or construct may still differ in measurement and application. T36–T38

illustrate this distinction: their average semantic similarity to the remaining tools is only $J = 0.21$, reflecting their different indicator provenance.

Pragmatic similarity displays a more polarized pattern (mean $J = 0.35$): 40.0% of pairs reach $J \geq 0.50$, yet 20.3% share no pragmatic coding. Identical combinations of intended users, spatial scale, and POS typology occur among 14.4% of pairs. Notably, T02 and T36–T38 share identical pragmatic profiles ($J = 1.00$) despite diverging at the empiric and semantic levels. Tools may therefore address similar users and settings while relying on non-equivalent measurement foundations and conceptual logics.

At the social level, 52.2% of pairs reach $J \geq 0.50$, and none return zero overlap. This convergence is partly produced by the uniform classification of all tools as research-based. Development pathways and geographic contexts nevertheless retain substantial differentiation, as shown by the lower similarity of T36–T38 to the wider set $J = 0.26$.

The Jaccard similarity analysis results show that convergence within one level does not predict convergence across the complete assessment system. Tools can share intended users, structural forms, or construct labels while diverging in how evidence is produced, organized and legitimized. The resulting comparability problem is therefore not indicator diversity alone but cross-level structural misalignment. Legitimate contextual variation becomes problematic when these relationships remain implicit. This finding provides the direct empirical basis for the semiotic compliance framework in Section 5.8, which seeks to make the architecture of each tool sufficiently explicit for responsible comparison, selection, and application.

5.8. Proposed Semiotic Compliance Framework

The preceding analyses demonstrate that divergence among POS assessment tools originates from differences extending beyond indicator selection to encompass the broader semiotic architecture through which assessment systems define, organize, interpret, and apply evidence. Collectively, the findings indicate that comparability is compromised when assumptions regarding measurement, structural organization, conceptual framing, intended application, and institutional legitimacy remain implicit or insufficiently documented. In response to these challenges, this study proposes a semiotic compliance framework that translates the empirical findings into a structured set of reporting and design principles intended to improve the transparency, internal coherence, and comparability of future POS assessment tools.

The framework does not seek to standardize public open space assessment by prescribing a universal set of indicators or measurement procedures. Such standardization would be neither feasible nor desirable because POSs differ substantially in physical characteristics, functional objectives, governance arrangements, cultural contexts, and user needs. Instead, the framework identifies the minimum information that assessment developers should explicitly document to enable users to understand what a tool measures, how it measures it, under which conditions it should be applied and how its outputs should be interpreted. Accordingly, semiotic compliance should be understood as transparency of assessment architecture rather than uniformity of assessment content.

At the empiric level, the analysis revealed considerable variation in operational specification despite widespread agreement regarding the general dimensions of POS quality. To improve measurement transparency, assessment tools should explicitly define each indicator, specify its measurement procedure, identify data sources, and describe the evidence required for assessment. For subjective indicators, survey questions, response scales and respondent groups should be clearly reported, while objective indicators should specify measurement protocols, spatial units, and scoring procedures. Such documentation reduces

ambiguity and enables meaningful interpretation and replication across different studies and application contexts.

At the syntactic level, the findings demonstrated substantial diversity in structural organization and procedural closure. Assessment developers should therefore clearly document the hierarchical structure of the tool, relationships among categories, subcategories, and indicators, methods of aggregation, weighting procedures and scoring algorithms. Equally important is the explicit clarification of whether the tool functions as a conceptual framework, semi-operational framework, or fully operational assessment instrument. Such transparency enables users to understand the level of procedural standardization provided by the assessment and the degree of flexibility available during implementation.

The semantic level addresses one of the principal sources of divergence identified in this study. Although many tools employ similar terminology, they frequently represent different conceptualizations of POS quality derived from distinct disciplinary traditions. Consequently, assessment tools should explicitly define the construct being measured, identify its theoretical foundations, and clarify the conceptual boundaries of the assessment. Where tools focus on specialized contexts, such as campus open spaces, neighborhood parks, or civic public spaces, these contextual distinctions should also be explicitly stated to prevent inappropriate comparison with instruments developed for fundamentally different purposes. Such clarification strengthens construct validity while reducing the risk of conceptual equivalence being inferred solely from shared terminology.

The pragmatic level emphasizes that assessment tools derive meaning from their intended use as much as from their measurement procedures. The analyses demonstrated substantial variation in intended users, assessment purposes, spatial scales, and application contexts. Accordingly, developers should specify who the intended users are, which decisions the tool is designed to support, and under what circumstances the assessment should be applied. Explicit documentation of application scope enables practitioners to determine whether an instrument is appropriate for planning, design evaluation, management, certification, post-occupancy assessment, or policy development, thereby improving the practical validity of assessment outcomes.

At the social level, the findings indicate that assessment tools are embedded within broader institutional, professional, and cultural environments that shape both their legitimacy and adoption. Semiotic compliance, therefore, requires explicit reporting of the tool's development context, geographic applicability, validation process, disciplinary origin, and institutional basis. Such information enables users to evaluate the extent to which findings may be transferred across governance systems, cultural settings, and planning traditions. Rather than assuming universal applicability, the framework encourages transparent acknowledgement of contextual limitations and institutional assumptions underpinning each assessment system.

Taken together, these five levels establish a coherent framework for improving the design, documentation, and interpretation of POS assessment tools. Importantly, the framework recognizes that diversity among assessment systems is both inevitable and, in many cases, desirable. Differences in environmental conditions, planning objectives, cultural expectations, and governance structures legitimately require different assessment approaches. Semiotic compliance, therefore, does not seek to eliminate contextual diversity; instead, it distinguishes legitimate contextual variation from undocumented structural inconsistency. When the assumptions operating at each semiotic level are explicitly documented, differences among tools become interpretable rather than problematic, allowing users to understand the reasons underlying divergent assessment outcomes.

The proposed framework has practical implications for several stakeholder groups. For researchers, it provides a structured basis for developing and comparing assessment

instruments while improving methodological transparency and facilitating cumulative knowledge development. For municipalities, urban planners, and designers, the framework supports more informed selection of assessment tools by encouraging evaluations of conceptual fit, procedural characteristics, and intended application rather than relying solely on the breadth of indicator coverage. For campus administrators and institutional managers, it offers guidance for selecting assessment approaches consistent with the distinctive governance structures and operational objectives of educational environments. More broadly, the framework provides journal reviewers, policymakers, and practitioners with a common reference for evaluating the completeness and transparency of emerging POS assessment methodologies.

Ultimately, the semiotic compliance framework extends beyond the evaluation of existing tools by providing a foundation for the next generation of POS assessment systems. By requiring explicit alignment among measurement procedures, structural organization, conceptual definitions, application contexts, and institutional legitimacy, the framework promotes assessment systems that are not only more transparent and internally coherent but also more comparable across diverse planning and research contexts. In doing so, it addresses the central challenge identified throughout this study: improving comparability not by reducing methodological diversity, but by making the semiotic architecture of assessment systems explicit. The framework, therefore, represents a practical mechanism for strengthening both the scientific rigor and operational usefulness of future public open space assessment tools.

6. Conclusions

This study examined divergence among public open space (POS) assessment tools by examining them as layered socio-technical sign systems. The findings demonstrate that the persistent difficulty in comparing public open space (POS) assessment tools cannot be explained adequately by differences in indicator selection alone. Examining thirty-eight assessment tools through a five-level semiotic framework reveals that divergence is generated by the cumulative interaction of measurement practices, structural organization, conceptual framing, intended application, and institutional context. The empirical evidence demonstrates that this divergence is both systematic and pervasive: 57.9% of tools lack complete operational definitions, 55.3% rely exclusively on subjective indicators, 55.3% function as assessment frameworks rather than fully operational tools, and 100% are research-based instruments. Collectively, these patterns show that comparability problems arise from cross-level structural misalignment rather than from indicator diversity alone. Although variations at any single level may appear manageable, their combined effects produce cross-level structural misalignment that fundamentally limits the interpretability and comparability of assessment outcomes. The findings, therefore, shift the explanation of divergence from the content of individual indicators to the broader architecture through which assessment systems represent, organize, and communicate the meaning of POS quality.

Viewing assessment tools as socio-technical sign systems provides a broader theoretical understanding of how assessment outputs acquire meaning and authority. The proposed semiotic perspective demonstrates that identical or closely related indicators cannot be assumed to produce equivalent assessments when they are embedded within different measurement procedures, hierarchical structures, construct definitions, decision contexts, or legitimacy pathways. Equally important, this study distinguishes between contextual diversity and structural inconsistency. Variations arising from differences in public space typology, disciplinary perspective, governance arrangements, or application context are both expected and often desirable, whereas divergence resulting from undocu-

mented assumptions and weak alignment across semiotic levels undermines transparency, reproducibility, and meaningful comparison.

Methodologically, this study extends existing comparative research by moving beyond indicator-centered analyses toward a multidimensional evaluation of assessment systems. Mapping twenty empirically derived attributes across the empiric, syntactic, semantic, pragmatic, and social levels provides a structured means of identifying where and why divergence occurs while preserving the contextual flexibility required for POS assessment. This analytical perspective enables researchers to evaluate assessment tools as coherent systems rather than isolated collections of indicators, thereby offering a more comprehensive basis for understanding similarities, differences, and sources of non-equivalence among existing instruments.

The proposed semiotic compliance framework translates these insights into a practical framework for improving future POS assessment tools. Rather than advocating a universal indicator set or standardized assessment procedure, it establishes minimum expectations for documenting the architecture of an assessment system so that its assumptions become explicit and its outputs can be interpreted appropriately. By emphasizing transparency, internal coherence, and explicit alignment across the five semiotic levels, the framework supports more informed tool development, evaluation, comparison, and selection while respecting the legitimate diversity required to address different public space contexts, planning objectives, and governance environments.

These contributions have practical significance for both research and practice. Researchers can use the framework to design more coherent assessment instruments and to compare existing tools based on their complete assessment architecture rather than indicator overlap alone. Practitioners, planners, designers, and public authorities can make more defensible decisions by selecting tools whose conceptual foundations, evidence types, and intended applications align with the objectives of a particular assessment rather than assuming that more comprehensive indicator sets necessarily produce more appropriate evaluations. In this way, the semiotic compliance framework encourages greater transparency in both the development and application of POS assessment methodologies.

This study should nevertheless be interpreted within the context of its scope. The analysis was limited to assessment tools documented in the accessible literature and therefore depended on the completeness of publicly available descriptions. The identified sample represents a substantial cross-section of contemporary POS assessment tools. Still, it does not encompass every existing instrument, particularly those developed outside the major academic databases or reported in languages other than English. As the field continues to evolve, additional tools may introduce new forms of divergence that extend the patterns identified here. Future research should therefore examine the application of the semiotic compliance framework across broader geographical and institutional contexts; evaluate its usefulness during the development of new assessment tools; and investigate whether greater semiotic compliance improves the consistency, transparency, and practical adoption of POS assessment systems.

Ultimately, this study advances the understanding of public open space assessment by demonstrating that meaningful comparability depends less on reducing diversity than on making the architecture of assessment explicit. When measurement procedures, structural organization, construct definitions, intended uses, and institutional contexts are aligned and transparently documented, differences among tools become interpretable rather than problematic. By reframing divergence as a question of cross-level coherence rather than indicator variation alone, this study provides both a conceptual explanation for the current fragmentation of POS assessment and a practical pathway toward more transparent, robust, and comparable assessment systems.

Author Contributions: Conceptualization, J.A.-Q. and M.R.J.; methodology, J.A.-Q.; validation, J.A.-Q. and M.R.J.; formal analysis, J.A.-Q. and M.R.J.; investigation, J.A.-Q. and M.R.J.; resources, J.A.-Q. and M.R.J.; data curation, M.R.J.; writing—original draft preparation, M.R.J.; writing—review and editing, J.A.-Q., M.R.J. and Z.A.; visualization, M.R.J.; supervision, J.A.-Q. and Z.A.; project administration, J.A.-Q. and Z.A.; funding acquisition, Z.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available because the associated research project is still ongoing and is subject to the funder's data-sharing restrictions.

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

References

1. Addas, A.; Maghrabi, A. Social evaluation of public open space services and their impact on well-being: A micro-scale assessment from a coastal university. *Sustainability* **2021**, *13*, 4372. [\[CrossRef\]](#)
2. Huang, Y.; Ye, L.; Chen, Y. Sustainable urban landscape quality: A user-perception framework for public space assessment and development. *Sustainability* **2025**, *17*, 3992. [\[CrossRef\]](#)
3. Islam, M.; Sarker, D.; Hasan, J.; Momtaz, Z. Public perceptions on urban open space and city livability in Barishal, Bangladesh. *Geol. Ecol. Landsc.* **2025**, *9*, 173–182. [\[CrossRef\]](#)
4. John, M.; Turaev, S.; Al-Dabet, S.; Abdulghafor, R. Multidimensional assessment of public space quality: A comprehensive framework across urban space typologies. *arXiv* **2025**, arXiv:2505.21555.
5. Praliya, S.; Garg, P. Public space quality evaluation: Prerequisite for public space management. *J. Public Space* **2019**, *4*, 93–126. [\[CrossRef\]](#)
6. Tawfik Mohamed, S.; Mandour, A.; Baker, H. A critical review of quality assessment tools for public spaces. *Eng. Res. J.* **2023**, *177*, 255–274. [\[CrossRef\]](#)
7. Greco, S.; Ishizaka, A.; Tasiou, M.; Torrisi, G. On the methodological framework of composite indices: A review of the issues of weighting, aggregation and robustness. *Soc. Indic. Res.* **2019**, *141*, 61–94. [\[CrossRef\]](#)
8. Al-Kofahi, Y.; Abdelqader, A.; Al-Qawasmi, J. Social sustainability metrics: A critical synthesis of selected studies across different urban settings. *Sustain. Cities Soc. Adv.* **2026**, *2*, 100037. [\[CrossRef\]](#)
9. Al-Qawasmi, J.; Yaro, A.I. Beyond indicator lists: Aligning place-based health assessment tools using a semiotic framework. *Cities Health* **2026**, 1–16. [\[CrossRef\]](#)
10. Al-Qawasmi, J.; Yaro, A.I.; Al-Mahdy, O.E.; Adenle, Y.A.; Ashour, Z. Evaluating place-based health: Coverage and consistency of indicators across assessment tools. *Urban Sci.* **2026**, *10*, 367. [\[CrossRef\]](#)
11. OECD. *Handbook on Constructing Composite Indicators: Methodology and User Guide*; OECD Publishing: Paris, France, 2008. [\[CrossRef\]](#)
12. Pineo, H.; Glonti, K.; Rutter, H.; Zimmermann, N.; Wilkinson, P.; Davies, M. Urban health indicator tools of the physical environment: A systematic review. *J. Urban Health* **2018**, *95*, 613–646. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Al-Qawasmi, J. Exploring indicators coverage practices in measuring urban quality of life. *Proc. Inst. Civ. Eng.-Urban Des. Plan.* **2019**, *172*, 26–40. [\[CrossRef\]](#)
14. Kumar, A. Systematic literature review (SLR). In *Meta-Analysis in Clinical Research: Principles and Procedures*; Springer Nature: Singapore, 2023; pp. 7–14. [\[CrossRef\]](#)
15. Bass, L.; Semetko, H.A. Content analysis. In *Research Methods in the Social Sciences: An A–Z of Key Concepts*; Oxford University Press: Oxford, UK, 2021; pp. 56–62. [\[CrossRef\]](#)
16. Stamper, R.K. Signs, information, norms and systems. In *Signs of Work: Semiosis and Information Processing in Organisations*; Holmqvist, B., Ed.; De Gruyter: Berlin, Germany, 1996; pp. 349–398.
17. Liu, K. *Semiotics in Information Systems Engineering*; Cambridge University Press: Cambridge, UK, 2000.
18. Taket, A.; Checkland, P.; Holwell, S. Information, systems and information systems-Making sense of the field. *J. Oper. Res. Soc.* **1999**, *50*, 554. [\[CrossRef\]](#)
19. Chandler, D. *Semiotics: The Basics*, 2nd ed.; Routledge: London, UK, 2007.

20. Eco, U. *A Theory of Semiotics*; Indiana University Press: Bloomington, IN, USA, 1976.
21. Morris, C.W. *Signs, Language and Behavior*; Prentice-Hall: New York, NY, USA, 1946.
22. De Saussure, F. *Course in General Linguistics*; Baskin, W., Translator; Philosophical Library: New York, NY, USA, 1959.
23. Ismail, A.S.; Zhaharin, E.N. Built form properties as sign and symbols of patron political ideology. *J. Kejuruter*. **2017**, *29*, 87–96. [[CrossRef](#)] [[PubMed](#)]
24. Marotta, A.; Spallone, R.; Turco, M.L.; Zich, U.; Vitali, M.; Marchis, E.; Pavignano, M. Visual images and language in architecture: Signifier semiotics and meaning semiotics. In Proceedings of the International and Interdisciplinary Conference IMMAGINI? Brixen, Italy, 27–28 November 2017; p. 964.
25. Nanni, A.; Bellentani, F. The meaning making of the built environment in the Fascist city: A semiotic approach. *Signs Soc.* **2018**, *6*, 379–411. [[CrossRef](#)]
26. Stamper, R.K. *Information in Business and Administrative Systems*; John Wiley & Sons: New York, NY, USA, 1973.
27. Bollen, K.A. *Structural Equations with Latent Variables*; John Wiley & Sons: New York, NY, USA, 1989.
28. Lazarsfeld, P.F. Evidence and inference in social research. *Daedalus* **1958**, *87*, 99–109.
29. Stamper, R.K.; Liu, K.; Hafkamp, M.; Ades, Y. Understanding the roles of signs and norms in organisations: A semiotic approach to information systems design. *Behav. Inf. Technol.* **2000**, *19*, 15–27. [[CrossRef](#)]
30. Strydom, W.; Puren, K. A participatory approach to public space design as informative for placemaking. *Chall. Mod. Technol.* **2013**, *4*, 33–40.
31. Switalski, M.; Torres, M.G.; Grêt-Regamey, A. The 3P's of place-making: Measuring placemaking through the latent components of person, procedures and place. *Landsc. Urban Plan.* **2023**, *238*, 104817. [[CrossRef](#)]
32. Santos Nouri, A.; Costa, J.P. Placemaking and climate change adaptation: New qualitative and quantitative considerations for the 'Place Diagram'. *J. Urban.* **2017**, *10*, 356–382. [[CrossRef](#)]
33. Skarzauskiene, A.; Maciuliene, M. Assessment of digital co-creation for public open spaces: Methodological guidelines. *Informatics* **2019**, *6*, 39. [[CrossRef](#)]
34. Addas, A. Enhanced public open spaces planning in Saudi Arabia to meet National Transformation Program goals. *Curr. Urban Stud.* **2020**, *8*, 184–204. [[CrossRef](#)]
35. Project for Public Spaces (PPS). What Is Placemaking? Available online: <https://www.pps.org/article/what-is-placemaking> (accessed on 10 February 2025).
36. Lorenzo, M.; Ríos-Rodríguez, M.L.; Suárez, E.; Hernández, B.; Rosales, C. Quality analysis and categorisation of public space: Development and validation of the PSCOQ. *Heliyon* **2023**, *9*, e13861. [[CrossRef](#)] [[PubMed](#)]
37. Knobel, P.; Dadvand, P.; Alonso, L.; Costa, L.; Español, M.; Maneja, R. Development of the urban green space quality assessment tool (RECITAL). *Urban For. Urban Green.* **2021**, *57*, 126895. [[CrossRef](#)]
38. Mela, A.; Tousi, E.; Varelidis, G. Assessing urban public space quality: A short questionnaire approach. *Urban Sci.* **2025**, *9*, 56. [[CrossRef](#)]
39. Uysal Bilge, F. A comparison of the qualifications of urban public open spaces and the evaluation of their activities: A pilot study in Ankara, Turkey. *Online J. Art. Des.* **2023**, *11*, 370–392.
40. Asfour, O.S.; Hossain, S.T. Assessing the impact of design quality attributes of public open spaces on users' satisfaction: Insights from a case study in Saudi Arabia. *Architecture* **2025**, *5*, 55. [[CrossRef](#)]
41. Imam, A.; Helmi, M.; Alkadi, A.; Hegazy, I. Exploring the quality of open public spaces in historic Jeddah. *ACE Archit. City Environ.* **2023**, *18*, 12123. [[CrossRef](#)]
42. Wijaya, N.; Christiani, D. Assessing publicness in urban green open spaces: Lessons from Jakarta, Indonesia. *J. Urban Manag.* **2025**, *15*, 360–374. [[CrossRef](#)]
43. Aydin, D.; Ter, U. Outdoor space quality: Case study of a university campus plaza. *ArchNet-IJAR Int. J. Archit. Res.* **2008**, *2*, 189–203.
44. Salama, A.M. Design intentions and users' responses: Assessing outdoor spaces of Qatar University campus. *Open House Int.* **2009**, *34*, 82–96. [[CrossRef](#)]
45. Khassawneh, O.; Khasawneh, Y. Place-making attributes and patterns of use in the central campus outdoor plaza in the Hashemite University of Jordan. *Open House Int.* **2023**, *48*, 47–68. [[CrossRef](#)]
46. Faragallah, A. The impact of productive open spaces on urban sustainability: The case of El Mansheya Square, Alexandria. *Alex. Eng. J.* **2018**, *57*, 3695–3703. [[CrossRef](#)]
47. Abbasi, A.; Chow, A.T.; Bryce, K. Open space quality in deprived urban areas: User perspective and use pattern. *Landsc. Urban Plan.* **2016**, *153*, 98–109.
48. Soliman, A.; Hassan, G.F.; El-Husseiny, M.; Haggag, M. Design characteristics of university campus outdoor spaces and their stress-reducing quality as perceived by students: Investigation in the Egyptian context. *Buildings* **2024**, *14*, 631. [[CrossRef](#)]
49. Zohra, H.F. Planning and designing livable recreational open spaces in developing countries: Insights from Souk Ahras City, Algeria. *Cities* **2024**, *147*, 104802. [[CrossRef](#)]

50. Ernawati, J.; Adhitama, M.S.; Alhad, M.A. The changes in public open space usage and perceptual urban design qualities after the COVID-19 pandemic. *Proc. IOP Conf. Ser. Earth Environ. Sci.* **2024**, *1359*, 012040.
51. Haddad, M. Investigating the quality of open educational spaces: The case of Al-Bayt University Al-Mafraq/Jordan. *Asian J. Environ.-Behav. Stud.* **2013**, *4*, 11–22.
52. Abu-Ghazalah, S.; Al-Goussous, J. Quality of space and its relation to the social behavior in academic open space. *Habitat Int.* **2009**, *34*, 278–286.
53. Addas, A.; Maghrabi, A. A proposed planning concept for public open space provision in Saudi Arabia: A study of three Saudi cities. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5970. [[CrossRef](#)] [[PubMed](#)]
54. Jamhawi, M.D.; Al-Shawabkeh, R.K.; Alobaidat, E.I. Spatial modeling of transformation of public open spaces in Zarqa, Jordan. *Landsc. Urban Plan.* **2020**, *198*, 103777. [[CrossRef](#)]
55. Thilakaratne, R. Designing liveable urban open spaces in high density cities. *Proc. IOP Conf. Ser. Earth Environ. Sci.* **2019**, *264*, 012013.
56. Kozlova, L.; Kozlov, V. Ten quality criteria of the public spaces in a large city. *Proc. IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *451*, 012101.
57. Rad, V.B.; Ngah, I.B. Assessment of quality of public urban spaces. *J. Appl. Sci.* **2014**, *14*, 1008–1015.
58. Ahirrao, P.; Khan, S. Assessing public open spaces: A case of city Nagpur, India. *Curr. Urban Stud.* **2021**, *9*, 478–500. [[CrossRef](#)]
59. Diaz-Sarachaga, J.M.; Jato-Espino, D.; Castro-Fresno, D. Evaluation of LEED for Neighbourhood Development and Envision Rating Frameworks for Their Implementation in Poorer Countries. *Sustainability* **2018**, *10*, 492. [[CrossRef](#)]
60. Salati, M.; Bragança, L.; Mateus, R. Sustainability Assessment on an Urban Scale: Context, Challenges, and Most Relevant Indicators. *Appl. Syst. Innov.* **2022**, *5*, 41. [[CrossRef](#)]
61. Hajr, M.; AbouElSeoud, T.; Hesham, B.; Qotp, S. The suitability of the MOSTADAM system with the global systems of sustainability in the desert city, case study (North Narjes project in KSA). *IOP Conf. Ser. Earth Environ. Sci.* **2024**, *1283*, 012015. [[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.