

UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

DIPARTIMENTO DI INGEGNERIA CIVILE, ARCHITETTURA, TERRITORIO, AMBIENTE E MATEMATICA

DOTTORATO DI RICERCA IN INGEGNERIA CIVILE, AMBIENTALE, DELLA COOPERAZIONE INTERNAZIONALE E DI MATEMATICA

ICAR/05

CICLO XXXVIII

Quality as a Driver of Sustainable Multimodal Mobility: From an Assessment Framework to Policy Applications in Urban Contexts

DOTT.SSA LAURA FERRETTO

PROF. BENEDETTO BARABINO

ARCH. MARTINA CARRA, Università degli Studi di Brescia

ABSTRACT

The increasing complexity of urban mobility systems and the shift toward sustainable multimodal transport require assessment frameworks capable of representing the complexity of quality in a context where multiple transport systems operate and interact with each other (i.e., non-motorised mobility, e-micromobility, Local Public Transport). Despite extensive research on transport quality, existing approaches remain fragmented, predominantly mode-specific, and rarely address the full travel chain. This PhD thesis responds to this gap by developing a comprehensive Multimodal Quality Assessment Framework that combines objective measurements, subjective evaluations, and a structured quality gaps management process to support evidence-based mobility planning.

The research begins with a systematic review of the literature, identifying conceptual fragmentation and the absence of a unified multimodal perspective. Building on this foundation, a new definition of multimodal transport quality is proposed and operationalised through the Logical Framework Approach (LFA).

A first rigorous methodological phase is dedicated to the selection of Key Quality Parameters (KQPs), combining expert judgements, AHP-derived weighting, Monte Carlo simulations, and citation/bibliometric validation. This integrated process leads to a robust and transferable set of KQPs suitable for diverse urban contexts.

A second methodological phase focuses on the measurement of these parameters through both objective and subjective measurements. Objective measurements are based on the Mystery Client (MC), while subjective measurements are captured through the Customer Satisfaction (CS). MC and CS represent the two tools used to carry out objective and subjective quality measurements, referring specifically to the four quality dimensions: designed, delivered, perceived, and expected.

Finally, a third methodological phase addresses the management dimension, identifying quality gaps across the four quality dimensions, and translating these gaps into targeted preventive or corrective measures.

The quality assessment framework is applied to the case study of Verona, where multimodal quality is assessed through both MC field measurements and Customer Satisfaction CS surveys. The integrated evaluation of designed, delivered, perceived, and expected quality enables a detailed Quality Gap Analysis, revealing substantial misalignments across criteria and the analysed O-D pairs. The findings support the formulation of targeted preventive and corrective measures aimed at improving accessibility, availability, comfort, customer care, environmental impact, information, safety & security, and time.

Overall, the thesis advances the scientific understanding of multimodal quality and offers a transferable, policy-relevant tool for guiding strategic decisions in sustainable urban

mobility planning. The framework provides public authorities and operators with a structured, user-centred approach to diagnosing and monitoring system quality and designing effective interventions, contributing to more efficient and sustainable multimodal mobility systems.

ABSTRACT

La crescente complessità dei sistemi di mobilità urbana e la transizione verso un sistema di trasporto multimodale e sostenibile richiedono quadri di valutazione capaci di rappresentare in modo sistemico la qualità in un contesto in cui molteplici sistemi di trasporto operano e interagiscono tra loro (ad esempio, la mobilità non motorizzata, la micromobilità elettrica e il Trasporto Pubblico Locale). Sebbene la qualità dei trasporti sia stata ampiamente indagata in letteratura, gli approcci esistenti risultano frammentati, prevalentemente specifici per modalità e solo raramente in grado di considerare l'intera esperienza di viaggio. La presente tesi di dottorato colma tale lacuna attraverso lo sviluppo di un quadro completo di valutazione della qualità multimodale, che integra misurazioni oggettive, valutazioni soggettive e un processo strutturato di gestione delle lacune di qualità, con l'obiettivo di supportare una pianificazione della mobilità basata sull'evidenza.

La ricerca prende avvio da una revisione sistematica della letteratura, che evidenzia la frammentazione concettuale esistente e l'assenza di una prospettiva multimodale unificata. Su queste basi, viene proposta e resa operativa una nuova definizione di qualità del trasporto multimodale mediante l'applicazione del *Logical Framework Approach* (LFA).

Una prima fase metodologica, rigorosamente strutturata, è dedicata alla selezione dei Parametri Chiave di Qualità (KQPs). Tale processo integra giudizi di esperti, ponderazioni derivate *dall'Analytic Hierarchy Process* (AHP), simulazioni Monte Carlo e validazione citazionale e bibliometrica, conducendo alla definizione di un insieme robusto e trasferibile di *KQPs*, adattabile a differenti contesti urbani.

La seconda fase metodologica è finalizzata alla misurazione dei *KQPs* attraverso un'integrazione di misurazioni oggettive e soggettive. Le misurazioni oggettive sono condotte mediante la tecnica del *Mystery Client* (MC), mentre le misurazioni soggettive sono rilevate tramite indagini di *Customer Satisfaction* (CS). *MC* e *CS* costituiscono i due strumenti adottati per valutare la qualità nelle sue quattro dimensioni fondamentali: progettata, erogata, percepita e attesa.

La terza fase metodologica affronta la dimensione gestionale, identificando le lacune qualitative nelle quattro dimensioni della qualità e traducendole in misure preventive e correttive mirate, volte a ridurre i divari riscontrati.

Il framework di valutazione della qualità è applicato al caso di studio di Verona, dove la qualità multimodale viene analizzata attraverso misurazioni sul campo mediante *MC* e tramite indagini di *CS*. La valutazione integrata della qualità progettata, erogata, percepita e attesa consente un'analisi approfondita del divario di qualità, evidenziando disallineamenti significativi tra criteri e coppie O-D analizzate. I risultati ottenuti supportano la definizione di misure preventive e correttive mirate al miglioramento

dell’accessibilità, della disponibilità, del comfort, dell’assistenza clienti, dell’impatto ambientale, dell’informazione, della sicurezza e del tempo.

Nel complesso, la tesi contribuisce all’avanzamento della comprensione scientifica della qualità multimodale e propone uno strumento trasferibile e rilevante per le politiche pubbliche, in grado di orientare le decisioni strategiche nell’ambito della pianificazione della mobilità urbana sostenibile. Il framework fornisce alle autorità pubbliche e agli operatori un approccio strutturato e incentrato sull’utente per diagnosticare, monitorare e migliorare la qualità del sistema, supportando la progettazione di interventi efficaci e promuovendo sistemi di mobilità multimodale più efficienti e sostenibili.

ACRONYMS

AASHTO: American Association of State Highway and Transportation Officials

AHP: Analytical Hierarchy Process

AMOS: Analysis of Moment Structures

ANOVA: Analysis of Variance

ANP: Analytic Network Process

AVE: Average Variance Extracted

AWKQPS: Adjusted Weight Key Quality Parameters

BSS: Bike-Sharing Systems

BLOS: Bicycle Level of Service

BWM: Best-Worst Method

CAPI: Computer-Assisted Personal Interviewing

CATI: Computer-assisted telephone interviews

CAWI: Computer Assisted Web Interviewing

CBD: Central Business District

CEN/TC: European Committee for Standardization/ Technical Committee

CFA: Confirmatory Factor Analysis

CFS: Contingent Field Survey

CS: Customer Satisfaction

DANP: DEMATEL-based Analytic Network Process

DBs: Databases

DEMATEL: DEcision MAking Trial and Evaluation Laboratory

EFA: Exploratory Factor Analysis

E-MM: Electric Micro-Mobility

E-PMVs: Electric Personal Mobility Vehicles

EQ: Expected Quality

Eqns: Equations

FHWA: Federal Highway Administration

GIS: Geographical Information System

GPS: Global Positioning System

HAN: Handling Algorithm

ICT: Information and Communication Technologies

ITS: Intelligent Transport Systems

IPA: Importance-Performance Analysis

KMO: Kaiser-Meyer-Olkin

KQPs: Key Quality Parameters

LFA: Logical Framework Approach

LINGO: LINear Generalized Optimization

LOS: Level of Service

LPT: Local Public Transport

MATLAB: MATrix LABoratory

MCDA: Multi-Criteria Decision Analysis

MC: Mystery Client

MEGA: Multi-faceted Evaluation of Grid Attributes

MIMIC: Multiple Indicators Multiple Causes

MMLOS: Multimodal Level of Service

MMQUAL: MicroMobility QUALity

MMTH: Multimodal Transport Hub

MNL: Multinomial Logit

NASEM: National Academies of Sciences, Engineering, and Medicine

N-MM: Non-motorized Mobility

NPE: Node-Place-Experience

O-D: Origin-Destination

OLS: Ordinary Least Squares

OPM: Ordered Probit Models

OUT: Output Algorithm

OVIs: Objectively Verifiable Indicators

PAs: Public Administrations

PCA: Principal Component Analysis

PEQUI: Pedestrian Environment Quality Index

PERS: Pedestrian Environment Review System

PLOS: Pedestrian Level of Service

POSet: Partially Ordered Set

PPT: Public Passenger Transport

PQ: Perceived Quality

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PSPNet: Pyramid Scene Parsing Network

PTCs: Public Transport Companies

P-TRANSQUAL: Public Transport Service Quality

PUC: Percentage of Users in Conformance

PUI: Percentage of Users for which a KQP is important

PUS: Percentage of Users Satisfied

PUT: Percentage of Users Targeted

QG: Quality Gap

QGA: Quality Gap Analysis

QoS: Quality of Service

Q-PLOS: Quality of Pedestrian Level of Service

SEM: Structural Equation Modelling

SERVPERF: Service Performance

SERVQUAL: ServiceQuality

SLOS: Scooter Level of Service

SPSS: Statistical Package for Social Science or Statistical Product and Service Solutions

SQ: Service Quality

SQI: Service Quality Index

STATA: Statistics and Data

TCQSM: Transit Capacity and Quality of Service Manual

USAID: United States Agency for International Development

VIKOR: VIseKriterijumska Optimizacija I Kompromisno Resenje [Multi-criteria Optimization and Compromise Solution]

WKQPS: Weight Key Quality Parameters

WoS: Web of Science

3FT: Three-Factor Theory

DEFINITIONS

Designed Quality // The level of service established by responsible organisations (e.g., public transport operators, shared micromobility providers, public authorities) through strategic and operational planning; it represents one of the four quality dimensions, together with delivered, expected, and perceived quality.

Delivered Quality // The actual level of service provided and observed in real operating conditions; it represents one of the four quality dimensions, together with designed, expected, and perceived quality.

Expected Quality // The level of service that users expect and consider satisfactory; it represents one of the four quality dimensions, together with designed, delivered, and perceived quality.

Multimodal transport // Multimodal transport refers to an urban mobility system in which a single trip is performed through the coordinated use of multiple transport modes, such as public transport, walking, cycling, and micromobility services. It relies on the physical, operational, informational, and fare integration of services and infrastructures to ensure a seamless, efficient, and user-centred travel experience. Multimodal transport aims to optimise accessibility, sustainability, and overall system quality by enabling travellers to combine modes flexibly according to trip needs and contextual conditions.

Quality Criteria // Quality Criteria are high-level dimensions that define the core domains of transport system quality. They represent the overarching conceptual categories used to evaluate whether a transport system is satisfactory in terms of quality, meeting user needs, and service standards. Examples include accessibility, comfort, safety and security, information, environmental impact, availability, customer care, and time. They serve as the broadest analytical layer, guiding the organisation and interpretation of quality parameters across all sustainable modes.

Quality Parameters // Quality Parameters are specific elements that operationalise each quality criterion and thematic clusters. They represent measurable features or attributes of a transport system that contribute to overall quality. Parameters define what is being assessed and form the core variables used to measure both perceived and delivered quality across modes.

Quality Indicators // Quality Indicators are the measurable units used to quantify each quality parameter during assessment. They provide a concrete method for observation and evaluation, typically expressed through numerical measures, scales, or ratings. Indicators define how a parameter is measured and enable objective monitoring of quality.

Perceived Quality // The users' subjective assessment of the service based on their direct experience; it represents one of the four quality dimensions, together with designed, delivered, and expected quality.

Sustainable mobility // Sustainable mobility refers to a mobility system that meets current transport needs while minimising negative environmental, social, and economic impacts and preserving the ability of future generations to meet their own mobility needs. It promotes efficient, low-carbon, and inclusive transport solutions that reduce resource consumption, emissions, congestion, and land use, while improving accessibility, safety, public health, and quality of life. Sustainable mobility prioritises collective and active transport modes, shared mobility services, and the integration of multimodal systems to support equitable and resilient urban development.

Urban environment // Urban environment refers to the dense physical, social, and functional context of cities and metropolitan areas, characterised by a high concentration of population, services, economic activities, and infrastructure. It includes the built environment, public spaces, and interconnected transport networks that support daily life. Urban environments typically generate higher and more complex mobility needs, where multimodal transport options and user experience play a central role in shaping travel behaviour.

Users // Users are individuals who interact with and utilise urban transport systems, including public transport, active mobility modes (such as walking and cycling), micromobility services, and intermodal facilities. They encompass both current and potential travellers, with diverse needs, behaviours, capabilities, and expectations. Users may also include vulnerable groups (such as children, older adults, and persons with reduced mobility) whose accessibility and safety requirements must be considered in transport planning and service design.

TABEL OF CONTENTS

1. INTRODUCTION..... 1
1.1 Motivations, research questions and objective..... 2
1.2 Scope and boundaries of the analysis..... 3
1.2.1 Analytical boundaries..... 3
1.2.2 Spatial and geographical boundaries..... 4
1.2.3 Modal boundaries..... 4
1.2.4 Temporal boundaries..... 5
1.2.5 Conceptual boundaries..... 5
1.2.6 Data boundaries..... 5
1.3 Structure of the thesis..... 6
2. SYSTEMATIC LITERATURE REVIEW..... 9
2.1 Methodological framework: Application of the PRISMA Statement..... 10
2.1.1 Identification..... 10
2.1.2 Screening..... 12
2.1.3 Eligibility..... 14
2.1.4 Inclusion..... 15
2.2 Results of the systematic literature review..... 15
2.2.1 Mapping the scientific community on transport quality..... 17
2.2.2 Evolution of the quality concept in transport systems..... 21
2.2.3 Rationale and purpose of transport quality assessment..... 27
2.2.4 Methods and techniques for measuring quality..... 38
2.2.5 Quality criteria and parameters in the literature..... 42
2.3 Discussion and identification of research gaps..... 47
3. THE NEW CONCEPT OF MULTIMODAL TRANSPORT QUALITY..... 49
3.1 Dimensions and framework for assessing transport quality..... 50
4. METHODOLOGY..... 55
4.1 Characterization of transport quality and Selection of Key Quality Parameters. 57
4.1.1 Identification of KQPs..... 59
4.1.2 Expert-based evaluation of KQPs..... 61
4.1.3 Assessment framework for KQPs..... 65
4.2 Measurement of quality in multimodal transport systems..... 67
4.2.1 Objective measurements..... 69
4.2.1.1 Preliminary Assessment Through a Simplified Delivered Quality Procedure 74
4.2.2 Subjective measurements..... 75
4.3 Management of quality in multimodal transport systems..... 78
4.3.1 Quality Gaps Analysis..... 79
4.3.2 Definition of improvement strategies..... 81
5. RESULTS AND EMPIRICAL APPLICATIONS..... 83
5.1 Characterization of quality and Selection of KQPs..... 84

5.1.1 Identification of KQPs from the literature..... 84
5.1.2 Expert judgment and evaluation of KQPs..... 87
5.1.3 Assessment and prioritization of KQPs..... 90
5.1.4 Discussion and methodological implications..... 93
5.2 Quality measurement in multimodal transport systems..... 97
5.2.1 Case study I: Melbourne, Australia..... 97
5.2.2 Case study II: Verona, Italy..... 112
5.2.2.1 Customer Satisfaction (CS) 113
5.2.2.2 Mystery Client (MC) 130
5.3 Quality management in multimodal transport systems..... 138
5.3.1 Quality Gaps Analysis..... 138
5.3.2 Definition of improvement strategies..... 147
5.3.2.1 Preventive Measure..... 147
5.3.2.2 Corrective Measure..... 152
5.3.3 Stakeholder roles in the implementation of quality measures..... 159
5.3.4 Concluding Remarks..... 161
6. FINAL DISCUSSION & CONCLUSIONS..... 163
6.1 Advancing Quality Assessment Through an Integrated Framework..... 164
6.2 Recommendations for practitioners and policy-makers..... 167
6.3 Limitations and directions for future research..... 167
6.4 Final Conclusions..... 168
7. REFERENCES..... 171
7.1 Publications arising from this research..... 179
8. APPENDIX..... 181
8.1 Appendix A..... 181
8.2 Appendix B..... 182
8.3 Appendix C..... 183
8.4 Appendix D..... 184
8.5 Appendix E..... 189

1. INTRODUCTION

Urban mobility systems are undergoing a deep transformation driven by the urgent need to enhance sustainability and integration across transport modes. Cities are increasingly required to address the growing complexity of travel behaviour, the diversification of mobility options, and the environmental and social challenges associated with urban transport (Banister, 2008). In this context, the quality of transport systems has emerged as a key determinant of both system performance and user satisfaction. Quality, however, remains a multidimensional and evolving concept that must be carefully defined, measured, and managed (Eboli & Mazzulla, 2010).

The transition towards sustainable and multimodal mobility requires new frameworks capable of assessing and managing quality not only within individual transport modes, such as local public transport (LPT) or non-motorized modes (N-MM), but also across their interactions. The growing presence of micro-mobility systems, such as shared e-scooters and other electric personal mobility vehicles (EPMVs), adds further complexity to the urban mobility ecosystem. Integrating these modes into a coherent, user-centred system is a major challenge for transport planners, operators, and policy-makers.

Despite the increasing interest in quality-based management approaches, existing methodologies often address quality from a sectoral or mode-specific perspective, failing to fully capture the multimodal and systemic nature of urban mobility (Moodley & Venter, 2022). Moreover, the relationship between expected, designed, delivered, and perceived quality remains underexplored, especially regarding how these dimensions influence user satisfaction and operational performance. The ability to evaluate and manage quality in a comprehensive, data-driven, and comparative manner, therefore, represents a significant research and policy gap.

This PhD research aims to address this gap by developing and testing a new framework for assessing and managing quality in sustainable, multimodal urban transport systems. The proposed framework integrates systematic literature analysis, expert-based evaluation, multicriteria decision analysis (MCDA), and empirical case studies to identify, measure, and manage the Key Quality Parameters (KQPs) that characterize the quality and user experience of diverse transport modes. By combining objective and subjective data sources, ranging from quality measurable indicators to user perceptions gathered through field investigations, this research contributes to a more holistic understanding of transport system quality.

Moreover, the study explores how quality assessment and quality gap analysis can support policy design and operational management, providing decision-makers with evidence-based tools to prioritize interventions and improve the integration of mobility services and infrastructure. Ultimately, the research seeks to enhance the sustainability

and integration of urban mobility systems through better quality governance and data-driven assessment methods.

1.1 Motivations, research questions and objectives

The overall objective of this research is to develop a comprehensive methodological framework for assessing and managing the quality of sustainable and multimodal urban transport systems. This framework is designed to integrate theoretical, methodological, and empirical components to bridge the gap between academic research and practical implementation.

To achieve this goal, the study is guided by four interrelated domains: conceptual, methodological, operational, and management.

1. Conceptual domain

- Which dimensions best represent the quality of different urban transport modes, including LPT, non-motorized modes, and micro-mobility systems?
- How can the concept of transport quality be redefined to reflect the challenges of sustainability, multimodality, and user-centred mobility?
- What are the key phases required to structure an effective framework for assessing the quality of multimodal transport systems?

2. Methodological domain

- How can most well-known methodological approaches, including multicriteria analysis and simulation techniques, contribute to the identification and selection of KQPs?
- What are the most appropriate techniques for integrating objective and subjective measurements of quality?
- How can quality gap analysis support the development of guidelines and strategies for improving the management and governance of multimodal transport systems?

3. Operational domain

- How can the proposed framework be applied to real urban contexts to measure and compare the quality of multimodal transport systems?
- What insights emerge from the empirical evaluation of different case studies in terms of strengths, weaknesses, and transferability?

4. Management domain

- In what ways can KQP-based evaluation contribute to improving mobility design, prioritizing interventions, and promoting multimodal integration?

- How can the outcomes of quality assessment be integrated into planning and management frameworks to support decision-making in sustainable urban mobility?

The thesis follows a progressive sequence of objectives, moving from conceptual reconstruction to methodological development and empirical validation. Specifically, it aims to:

1. Reconstruct the evolution of the concept of transport quality through a systematic literature review and define a new conceptual framework for assessing and managing the quality of multimodal transport systems.
2. Translate the conceptual framework into a methodological one by developing an integrated approach for:

- the identification, selection, and weighting of KQPs, combining expert judgment, multi-criteria decision analysis, and Monte Carlo simulation;
- the integrated measurement of the four dimensions of quality (designed, expected, delivered, perceived) through the combined use of objective and subjective approaches (Mystery Client and Customer Satisfaction techniques);
- the quality management based on the quality gap analysis and definition of guidelines to translate empirical findings into strategic and operational recommendations that enhance the integration, efficiency, and sustainability of multimodal urban mobility systems.

3. Apply and test the proposed methodology to selected urban case studies (Melbourne, Australia, and Verona, Italy).

Through these objectives, the thesis provides a replicable framework for assessing, measuring, and managing the quality of sustainable and multimodal transport systems in urban contexts.

1.2 Scope and boundaries of the analysis

This research is positioned within the broader domain of sustainable urban mobility, with a specific focus on the assessment of multimodal transport systems quality in urban contexts. The following boundaries - analytical, spatial, modal, temporal, conceptual, and data-related - together define the scope, framework's applicability, and the intended limitations of the research.

1.2.1 Analytical boundaries

The research adopts a system-level perspective, addressing the quality of urban mobility systems as an integrated whole rather than focusing on the design or optimization of individual modes or infrastructure components. Accordingly, the analysis does not cover aspects such as network optimization, infrastructure dimensioning, or demand

forecasting. Instead, it focuses on the assessment of service and infrastructure quality, as well as user experience, emphasizing how these dimensions interact within a multimodal and sustainable mobility ecosystem. The analytical framework integrates both operational and perceptual dimensions of quality, combining technical and user-centred approaches.

This dual perspective allows the study to capture not only how mobility systems perform, but also how they are experienced by users, which represents a fundamental element in the transition toward more sustainable travel behaviours.

1.2.2 Spatial and geographical boundaries

From a spatial standpoint, the study focuses on urban environments, where short- and medium-distance trips, multimodal interactions, and first- or last-mile connections are most prominent. Consequently, intercity and long-distance transport systems, including regional and high-speed rail, are excluded from the analysis, as they fall outside the functional boundaries of urban mobility.

The framework is designed to be transferable across different urban contexts, ranging from large metropolitan areas to medium-sized cities.

This rationale is reflected in the two case studies selected for empirical application:

- Melbourne (Australia) represents a large and complex metropolitan system, where surface metro and suburban rail services are part of the urban mobility network.
- Verona (Italy) represents a medium-sized European city without a metro system, characterized by surface LPT and growing integration with non-motorized and micromobility modes.

The comparative use of these two cities ensures that the proposed framework remains adaptable and robust across varying spatial scales, governance models, and infrastructural settings.

1.2.3 Modal boundaries

The study explicitly focuses on transport modes that contribute most directly to sustainable urban mobility, namely:

- LPT - including buses, trams, and light rail services- as the core of collective mobility systems within cities.
- Non-Motorized Modes (N-MM) - including walking and cycling, which are fundamental for promoting active, inclusive, and low-impact mobility.
- Electric Micromobility (E-MM) - including shared or individual small electric vehicles (e.g., e-scooters), which play a growing role in short urban trips and intermodal integration.

In contrast, private motorized transport, including electric cars, is excluded from the analysis.

Although electric vehicles are often framed as sustainable, they remain less efficient from a systemic and collective perspective compared to public and active modes.

The research thus aligns with the sustainability principle that collective and shared mobility represents a higher level of environmental and social efficiency than private vehicle use.

Within LPT, metro systems are given limited attention, though not completely excluded. This decision reflects both the representativeness of the selected case studies and the objective of ensuring broad applicability to cities where heavy rail infrastructure is not present. By focusing on modes typical of most urban contexts, the framework seeks to remain relevant and implementable beyond large metropolitan areas.

1.2.4 Temporal boundaries

The research adopts a cross-sectional analytical approach, relying on data collected during specific time periods within each case study. It does not include longitudinal analyses or predictive modelling of quality evolution over time. However, the methodological framework has been designed to remain applicable in future longitudinal or monitoring-oriented studies.

1.2.5 Conceptual boundaries

Conceptually, the study is rooted in the quality dimension of transport systems, understood as a multidimensional construct encompassing availability, accessibility, comfort, customer care, environmental impact, information, safety and security, and time.

While sustainability is an overarching principle, it is approached through the lens of quality assessment rather than through direct environmental impact analysis. In other words, the research assumes that improved service quality and user satisfaction act as drivers of sustainable mobility, promoting shifts toward more efficient and low-impact transport modes. The study also excludes aspects primarily related to economic or regulatory frameworks, such as cost-benefit analysis, fare policy, or governance structures, except where they directly affect service quality or user perception.

1.2.6 Data boundaries

The research relies on a combination of quantitative and qualitative data, including expert judgments, field observations, and users' surveys. It does not incorporate large-scale "big data" sources or real-time mobility data (e.g., from GPS, mobile apps, or sensors), as these would require different analytical infrastructures and objectives.

The data used are thus tailored to support methodological validation and comparative analysis, rather than system monitoring or predictive modelling.

1.3 Structure of the thesis

The thesis is structured into six main chapters, complemented by appendices and references, following a logical sequence that reflects the conceptual, methodological, and empirical progression of the research.

Chapter 1 – Introduction presents the overall context and rationale of the study, emphasizing the growing importance of quality as a key dimension of sustainable and multimodal urban mobility. It defines the research questions and objectives, outlines the scope and boundaries of the analysis, and provides a summary of the thesis structure.

Chapter 2 – Systematic Literature Review presents the findings of an analysis conducted in accordance with the PRISMA methodology. Through a structured and transparent process, the chapter outlines the scientific community engaged in the study of quality within sustainable transport systems; it explores the theoretical background and evolution of the concept of quality within the transport sector; it identifies the main rationales for assessing transport quality, the methodological approaches applied, and the quality criteria and parameters most frequently adopted in previous research.

Chapter 3 – The Concept of Transport Quality The chapter proposes a conceptual framework for interpreting transport quality in the context of sustainability and multimodality.

Chapter 4 – Methodology describes the methodological framework developed to assess, measure, and manage transport quality in multimodal and sustainable urban contexts. It is structured into three main phases: (i) the characterisation and selection of KQPs through expert judgement, multi-criteria decision analysis (AHP), and Monte Carlo simulation; (ii) the integrated measurement of transport quality using both objective (Mystery Client) and subjective (Customer Satisfaction) techniques; and (iii) the management of transport quality through gap analysis and formulation of strategic guidelines and recommendations.

Chapter 5 – Results and Empirical Applications report the outcomes of the methodology based on two urban case studies: Melbourne (Australia) and Verona (Italy). The Melbourne case study serves as an initial application of the proposed framework, focusing on the objective measurement of transport quality through the partial application of the Mystery Client technique. This partial implementation allows the identification of quality critical issues and the validation of the framework's operational components.

The Verona case study, instead, represents the complete empirical implementation of the methodology. It integrates both objective (Mystery Client) and subjective (Customer Satisfaction) assessments. Finally, it extends the quality measurements to the

management phase through the quality gap analysis and the formulation of strategic guidelines.

Chapter 6 – Final discussion & Conclusions interprets and critically analyses the results obtained, highlighting the methodological, conceptual, and operational implications of the research. It discusses the main quality gaps identified, their relevance for multimodal integration, and the potential of the proposed framework to support evidence-based mobility planning and management. Moreover, the chapter summarises the main findings of the thesis, underlining its scientific and practical contributions. It reflects on the methodological robustness and transferability of the proposed approach, outlines the main limitations, and suggests directions for future research and applications.

The thesis concludes with a comprehensive list of References (Chapter 7), followed by an Appendix (Chapter 8) containing supplementary data and materials.

Overall, the structure reflects a progressive research process: from conceptual foundations, through methodological development, to empirical validation and managerial implications. Ultimately, it provides a comprehensive and replicable framework for assessing and managing the quality of sustainable, multimodal urban mobility systems.

To illustrate this logical progression, *Figure 1* presents an overview of the research framework adopted in this thesis, summarising the three methodological phases

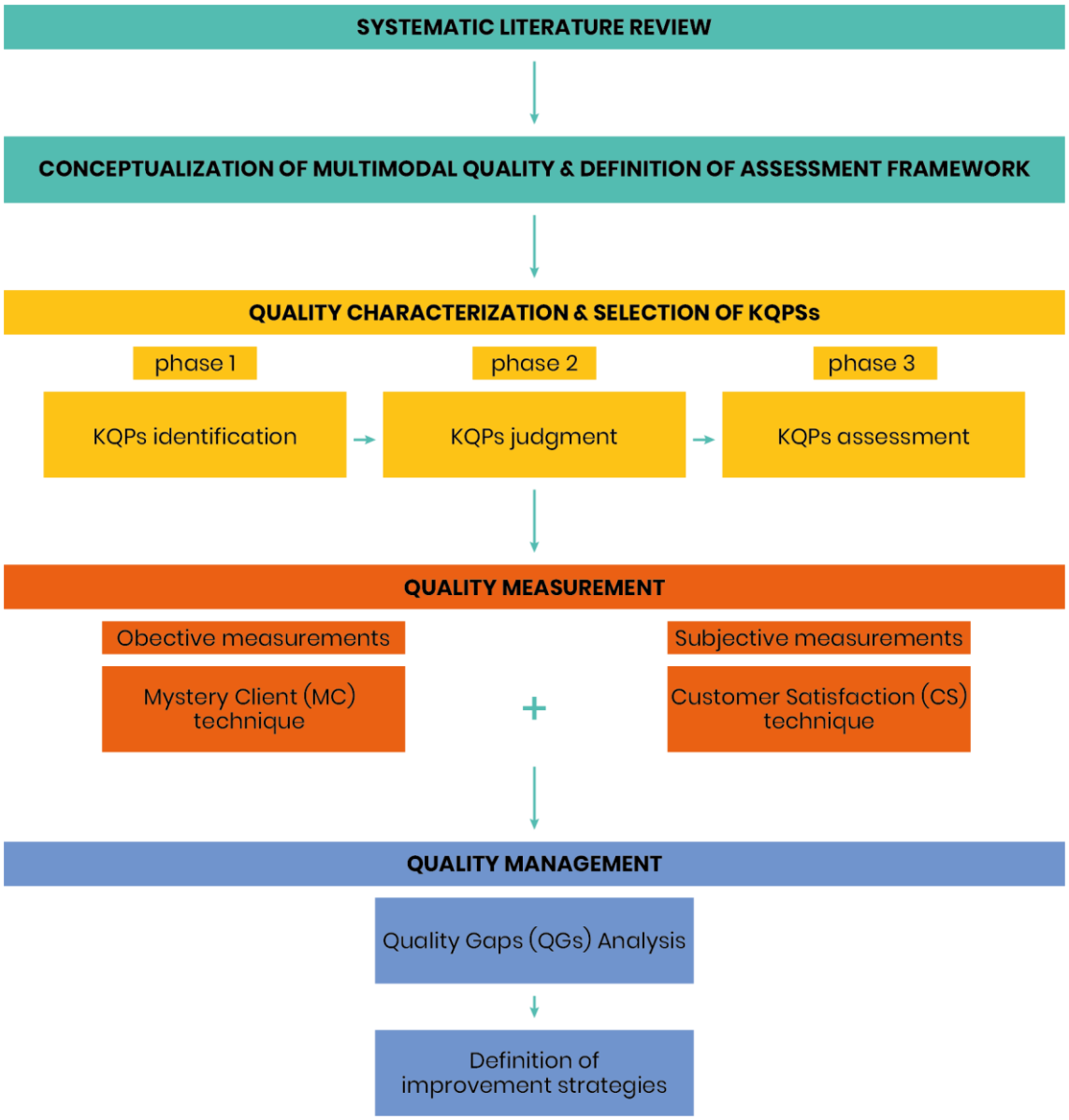


Figure 1. Overview of the research framework based on three main phases: (i) quality characterization and selection of KQPs, (ii) quality measurement, (iii) quality management.

2. SYSTEMATIC LITERATURE REVIEW

This chapter aims to systematize the existing body of knowledge on transport quality assessment in urban contexts. The objective is to identify, categorize, and critically analyse the quality thematic clusters, criteria, and parameters proposed and applied across different urban transport systems, ranging from LPT to non-motorized modes and electric micromobility.

Despite the growing academic and institutional interest in transport quality, the literature remains fragmented, with studies often limited to single transport modes or specific geographic contexts. Traditional reference frameworks, such as the Transit Capacity and Quality of Service Manual (Kittelson & Associates, 2013) and the European Standard EN 13816 (CEN/TC 320, 2002), have significantly contributed to defining quality in public transport services. However, they largely overlook the integrated and multimodal nature of contemporary urban mobility. Moreover, as new transport modes emerge, such as electric micromobility, the need for a comprehensive and comparable framework becomes increasingly evident.

To address this gap, a systematic literature review was conducted using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Moher et al., 2009). The PRISMA statement was originally developed to ensure methodological transparency and reproducibility in medical and social sciences; in this research, it is adapted to the transport domain to provide a rigorous, evidence-based synthesis of existing studies on quality assessment. This approach allows the identification of recurring quality criteria and parameters, the analysis of their methodological foundations, and the recognition of gaps that hinder the development of a unified assessment framework for multimodal transport quality in urban contexts.

The review outlines the state of knowledge about the quality of multimodal transport systems, investigating:

- the scientific community that contributes to the study of urban transport quality, and the main thematic areas;
- the underlying motivations of quality assessment, and how they relate to policy and management objectives;
- the methodological approaches and data sources to assess transport quality;
- the most frequently used quality criteria, parameters, and indicators across modes and contexts.

Through this systematic characterization, the chapter aims to provide both a descriptive overview of the field and a conceptual foundation for the subsequent phases of the research. The findings of the review not only inform the subsequent selection of KQPs but

also contribute to defining the multidimensional framework that supports the quality measurement and management approaches.

2.1 Methodological framework: Application of the PRISMA Statement

The systematic literature review was conducted following the PRISMA statement, which ensures transparency and reproducibility throughout the review process. In this research, the PRISMA was adapted to the field of sustainable urban mobility to systematically identify and synthesize scientific evidence on transport quality. The objective of this phase is to identify the key criteria, parameters, and methodological approaches used to assess the quality of sustainable and integrated urban transport systems.

As illustrated in *Figure 2*, the review process is articulated into four sequential stages: identification, screening, eligibility, and inclusion.

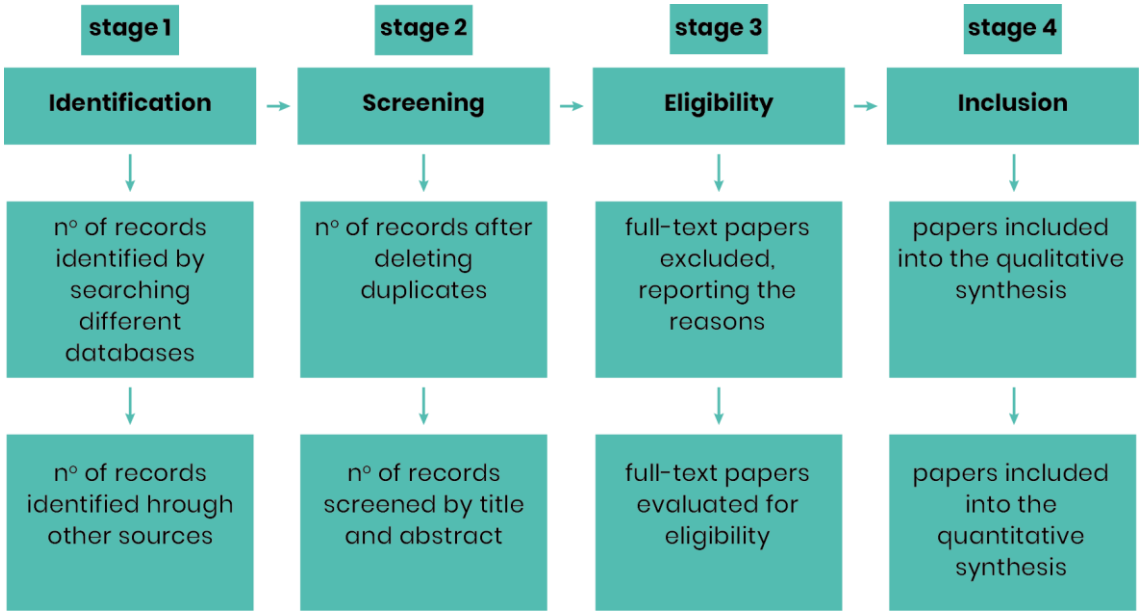


Figure 2. Flow chart of the systematic review process based on the PRISMA Statement.

2.1.1 Identification

The first methodological stage focused on identifying articles about the quality of urban transport systems, starting with a keyword search conducted in three databases (DBs): Web of Science (WoS), Scopus, and Google Scholar. These DBs supply high-quality documents and are among the best sources of bibliographic data (Baas et al., 2020). On Scopus and WoS the search was carried out on title, abstract and keywords, while on Google Scholar on title only, since the search system does not include the abstract and keywords. In all DBs, the same keywords were used to relate the quality assessment to

the forms of urban transport mentioned above. After obtaining the search results, a single DB containing all scientific papers and their characterization was structured.

Within the three chosen DBs, the keywords "quality" and "quality indicator" were cross-referenced with "assessing", "evaluating", "monitoring", and with "public transport", "collective transport", "sidewalks", "pedestrian mobility", "cycle lanes", "cycle paths", "cycle ways", "cycle routes", "bike sharing", "ePMVs", "electric personal mobility vehicles", "transport infrastructure", "pedestrian infrastructure", and "cycle infrastructure" for a total of 84 different combinations.

In the identification stage, searching for scientific papers through keywords resulted in 5,482 records: 144 on Google Scholar, 2,006 on WoS, and 3,332 on Scopus (*Figure 3*). The records on Google Scholar are markedly fewer because the search was not by abstract and keywords, since the system only allows searching by title, and searching for the overall text the number of papers and/or grey literature is too spread to be manageable.

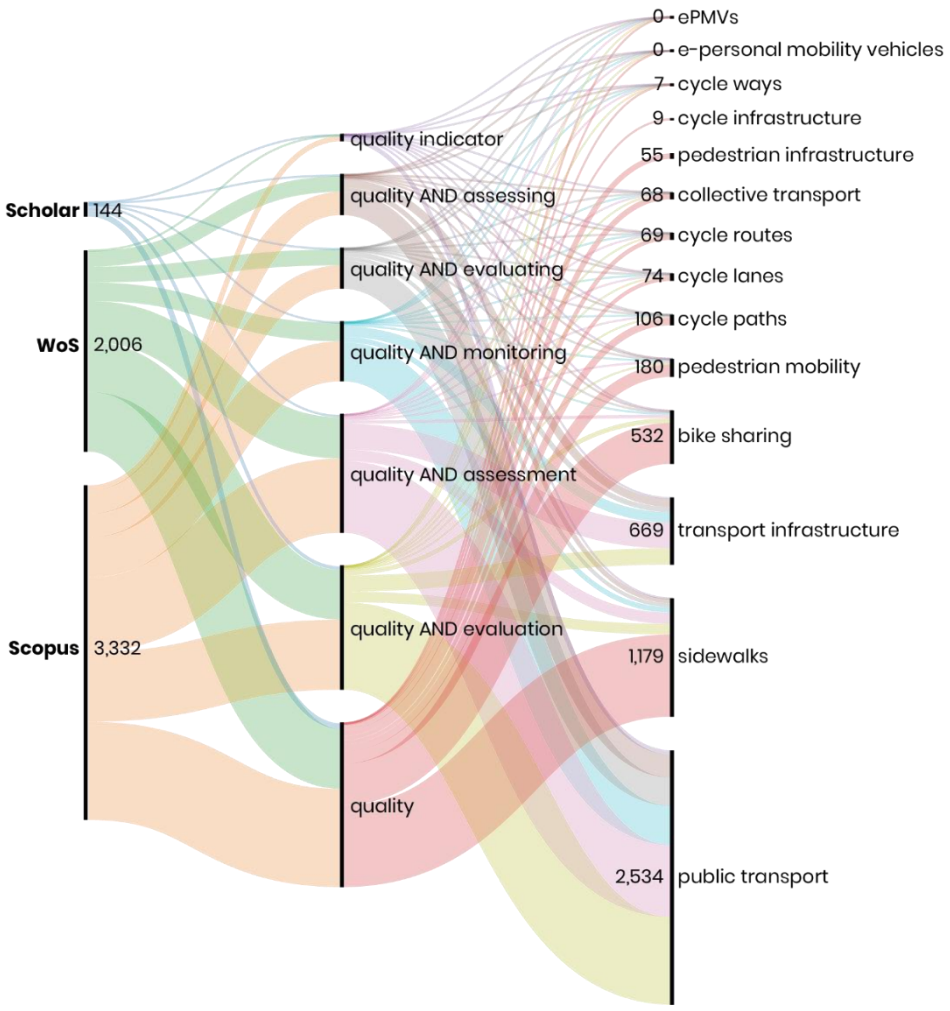


Figure 3. Summary diagram of the results obtained from the first stage of the PRISMA statement.

2.1.2 Screening

The second stage of screening is divided into three parts. The first part involved the elimination of duplicates within each DB and subsequently in all DBs once merged. This step led to a single DB containing the papers whose titles and abstracts contain one or more combinations of the chosen keywords. The DB is structured according to a matrix of rows corresponding to the papers and columns containing different information levels useful for characterizing scientific contributions. The second part is based on analysing the relevance of the titles of the papers presented in the overall DB concerning the research topic. Finally, the third part focuses on the abstract relevance of the remaining papers.

The screening stage guaranteed a restricted selection of papers and allowed the identification of more specific categories than those used in the identification stage. The obtained papers were then classified according to the different modes of transport to facilitate the subsequent phases of analysis and to identify the focus of the scientific contributions more precisely.

In the screening stage 1,516 double records were first removed in the individual DBs; subsequently, the three DBs were merged, and another 1,099 articles present in more than one of them were eliminated, thus going from 5,482 to 2,867. By analysing the title of each paper obtained from the previous step, 2,475 records were excluded, going from 2,867 to 392. Of these 392 papers, 153 were eliminated based on the abstract until 239 papers passed to the eligibility stage.

The clustering defined at the end of the screening stage reveals a distribution aligned with the one identified in the preceding "Identification" stage. The distribution of scientific papers, categorized by transport categories and modes and represented as semicircles or concentric arcs across four distinct groups (*Figure 4*), predominantly features studies on the quality of collective transport (specifically bus services), followed by research on non-motorized (pedestrian) mobility. Research on the quality of multimodal transport systems is notably limited, as is the case with studies concerning electric micro-mobility.

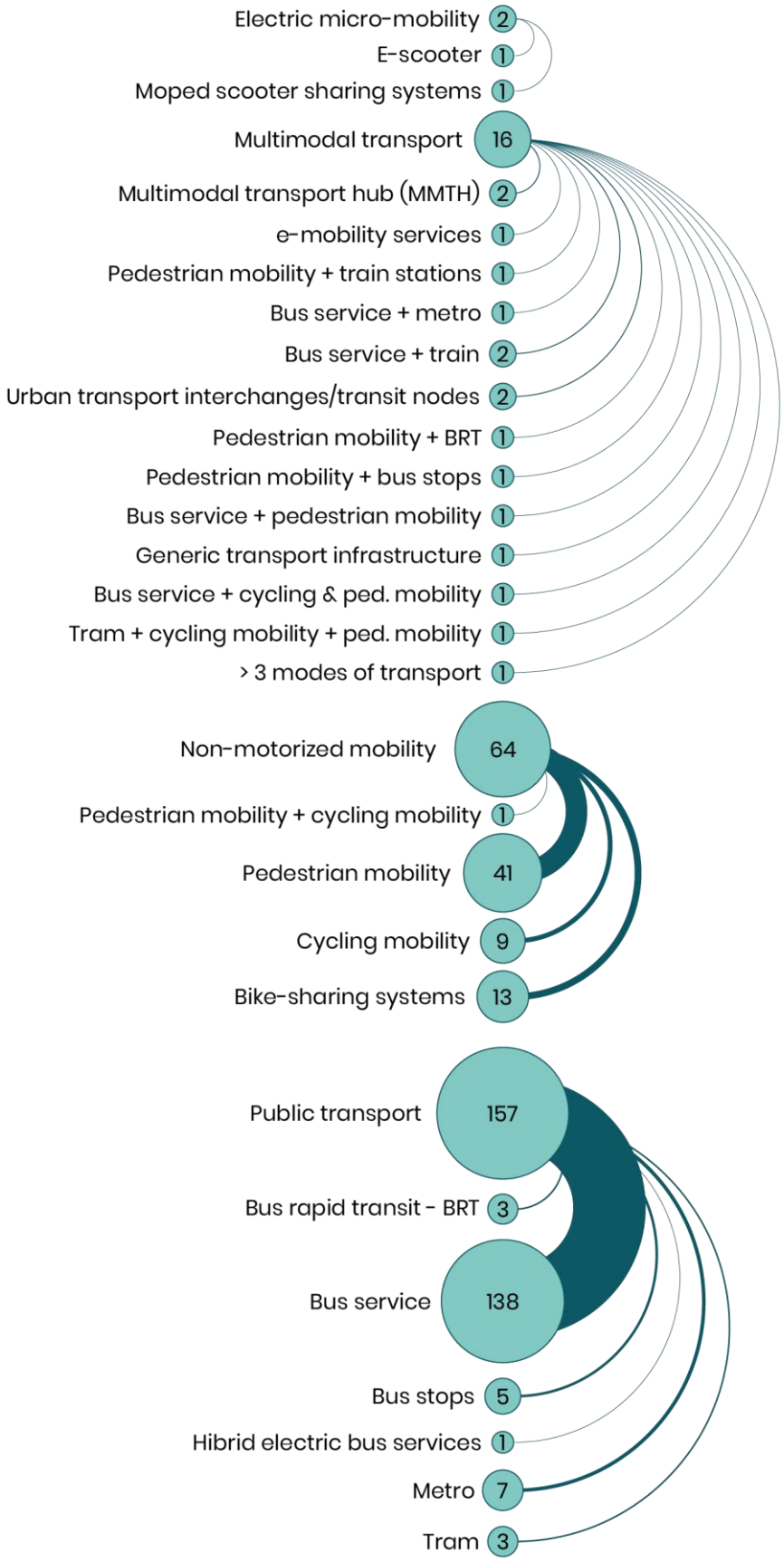


Figure 4. Number of articles divided by transport categories and transport modes.

2.1.3 Eligibility

The third stage of eligibility is based on the integral reading of the papers derived from the previous stage. Before proceeding to the full reading of the papers, those concerning bus services have been excluded, except those focusing on bus stops. This choice is justified by the fact that there are already paper reviews that provide a viable methodology for defining a set of quality criteria and parameters for evaluating LPT services (Barabino et al., 2020). Infrastructural rather than service aspects of LPT have been less investigated and are strictly integrated with the infrastructures dedicated to non-motorized mobility, especially pedestrian mobility. For this reason, studies on bus/train/metro stops have moved on to the eligibility stage. The same goes for studies focusing on forms of mass transit such as trams, which have been far less investigative than bus service. Papers on the metro were also excluded, as it is a transport system that is not common in all urban centres but is typical of large cities/metropolitan cities.

In light of the criteria defined after the screening stage, 90 papers have passed the eligibility stage. The integral reading of the papers led to further screening, identifying the scientific contributions that fully meet the criteria defined upstream of the development of the four stages of the PRISMA statement. At this stage, 23 papers were excluded because they did not fully meet the search criteria (*Table 1*) and 18 papers were excluded because they were not available and/or downloadable. Consequently, 49 papers passed to the final stage of inclusion. To complement the database search and ensure a more comprehensive coverage of the scientific literature, an ancestral search approach (backward snowballing) was also applied. This method consists of reviewing the reference lists of the papers included in the screening process to identify additional relevant studies that may not have been retrieved through Scopus, WoS, or Google Scholar. In several cases, key publications cited in the selected articles were not indexed in the three databases. These additional sources were therefore manually collected and assessed using the same inclusion and exclusion criteria applied in the main review. This complementary step strengthened the completeness of the literature sample and reduced the risk of missing influential contributions to the study of transport quality.

Table 1. Reasons for the exclusion of the records from the inclusion stage.

Reason for the papers' exclusion	**N° of papers excluded**
The article is not strictly related to the urban area.	1
The article is specific to an exceptional situation, such as a sporting event, which implies a change in temporary transport demand.	1
The article focuses on the point of view of the service provider rather than the users.	1
The article presents the same indicators as a further article by the same authors (so only one is considered).	1
The article is not in English.	2
The article provides general guidelines but does not include methods of evaluation of the quality of transport.	2
The article considers the transport quality only partially, and the evaluation was not coherent with the focus of this research.	15
The article is not available and/or downloadable.	18
Total	**41**

2.1.4 Inclusion

The inclusion stage represents the final step of the PRISMA process, in which all records that met the predefined eligibility criteria were consolidated into the final corpus of analysis. At this stage, 49 papers were retained, forming the evidence base for the subsequent synthesis and discussion. These publications collectively provide a comprehensive overview of the multidimensional nature of transport quality and its assessment within the urban context.

The selected studies encompass diverse transport modes, including pedestrian and cycling mobility, bike-sharing systems, electric micromobility, and LPT, thereby reflecting the multimodal perspective that guides this research. Their inclusion ensures that both infrastructure- and service-related aspects are represented, along with user-centered and operational dimensions of quality.

This stage also serves as a methodological transition from the screening and selection of relevant literature to the systematic analysis of its content. The consolidated corpus of 49 papers serves as the foundation for the next phase (Section 2.2), which examines the primary trends, methodological approaches, and conceptual dimensions emerging from the international scientific community. In particular, the following sections analyze the geographical and temporal distribution of research, the rationale behind quality assessment, the methods employed, and the KQPs and indicators identified in the literature.

2.2 Results of the systematic literature review

This section presents the main findings emerging from the systematic literature review, offering an integrated overview of how transport quality has been addressed within the scientific community. The analysis focuses on five complementary areas that together illustrate the current state of knowledge on quality assessment in sustainable and multimodal transport systems.

First, the mapping of the scientific community (Section 2.2.1) identifies the main contributors, publication outlets, and geographical distribution of research, highlighting how the topic has evolved over time and across disciplines.

Second, the evolution of the concept of quality (Section 2.2.2) in the transport sector traces how the notion of quality has changed over time and identifies conceptual gaps related to the multimodal dimension.

Third, the rationale and purpose of quality assessment (Section 2.2.3) explore the motivations driving the evaluation of transport system quality and its role in promoting sustainability, efficiency, and user satisfaction.

Fourth, the methods and techniques for measuring quality (Section 2.2.4) examine the methodological approaches applied in the literature, from traditional statistical analyses to advanced multi-criteria and computational frameworks.

Finally, the quality criteria and parameters (Section 2.2.5) synthesize the key elements used to define, structure, and quantify the quality of multimodal transport systems in urban contexts.

Together, these analyses form the empirical foundation for the conceptual and methodological framework developed in the subsequent phases of this research.

A preliminary overview of the research landscape can be derived from the keywords associated with the reviewed papers.

Figure 5 presents the keyword cloud generated from the 49 included studies, illustrating the dominant themes, transport modes, and methodological approaches that define current research on transport quality. The size of each term reflects its frequency of occurrence: the larger the word, the more frequently it appeared as a keyword in the analysed papers.

The keyword distribution highlights recurring concepts such as customer satisfaction, service quality, and perceived value, confirming the strong user-oriented perspective of most studies. Terms referring to bike-sharing, public bike systems, and micromobility indicate growing attention toward sustainable, shared, and active transport modes.

Methodologically, the prominence of multi-criteria decision-making tools (e.g., Best-Worst Method, VIKOR, fuzzy-set approaches) reflects the need to manage complex interdependencies among quality criteria. Statistical techniques such as Structural Equation Modelling (SEM), factor analysis, and latent variable modelling further demonstrate the field's methodological rigor in capturing perceptions and service quality.

Spatial and design-related terms (e.g., GIS, walkability, Pedestrian Level of Service, urban design qualities) emphasize the increasing integration of geospatial and urban planning approaches in assessing pedestrian environments. Keywords such as disability, age-sensitive design, and elderly people point to the growing inclusion of equity and accessibility dimensions.

Emerging concepts like smart mobility, machine learning, and deep learning underline the progressive adoption of data-driven and computational methods in transport quality evaluation.
Overall, the keyword profile reveals a comprehensive research field that combines user-centred evaluation, advanced quantitative methodologies, spatial analysis, and inclusive design principles to support sustainable and high-quality urban mobility systems.


quality
pedestrian
service
transport
walkability
bike-sharing
design
urban
perception
analysis
public
pedestrians
mobility
accessibility
sustainable
walking
sidewalk
safety
fuzzy
factor
value
monitoring
environment
sidewalks
infrastructure
comfort
level
satisfaction
model
criteria
index
vikor
gis
bicycle
planning
elderly
stops
system
transportation
evaluation
qualities
street
tourism
bus
indicators
low-quality
road
travel
walkability
transit
user
facilities
cities
learning
assessment
physical
experience
decision
active
mixed
variables
bwm
multi
sem
valuations
social
recreational
components
latent
ordinary
ordered
route
environmental
set
collaborative
charts
activity
trip
review
bike
anthropometry
multi-stakeholders
effectiveness
modeling
acceptance
framework
lane
sensitive
neighbourhoods
attribute
customer
station
configuration
hub
multilevel
principal
inspections
streetscape
non-aggregative
confirmatory
preference
masterton
old
spatial
rating
ranging
protection
facility
perceived
seniors
scaling
nodes
mid-sized
cluster
university
standards
cycling
choice
mcdm
experience
image
detection
appearance
buses
metro
three-factor
sensory
secret
stop
land
ped cycle
condition
instrumented
ahp
services
accessible
discrete
nuisance
shopper
control
behaviour
grey-vikor
importance
grey-danp
dimensions
methodology
istanbul
mmqual
economy
dimensions
importance-performance
network
dunok
issues
walkway
young
modes
cognition
tram
transfer
people
security
campus
vehicles
integration
stated
municipal trolley
space
regression
machine
granada
automatic
attractiveness
micromobility
online
asset
multimodal
c-means
methodology
partially
method
theory
equation
disability
built
city
plos
multi-criteria
perception
integration
node-place
best-worst
non-aggregative
confirmatory
path
agg
trips
inspections
principal
multilevel
choice
mcdm
image
valuations
ordinary
latent
detection
components

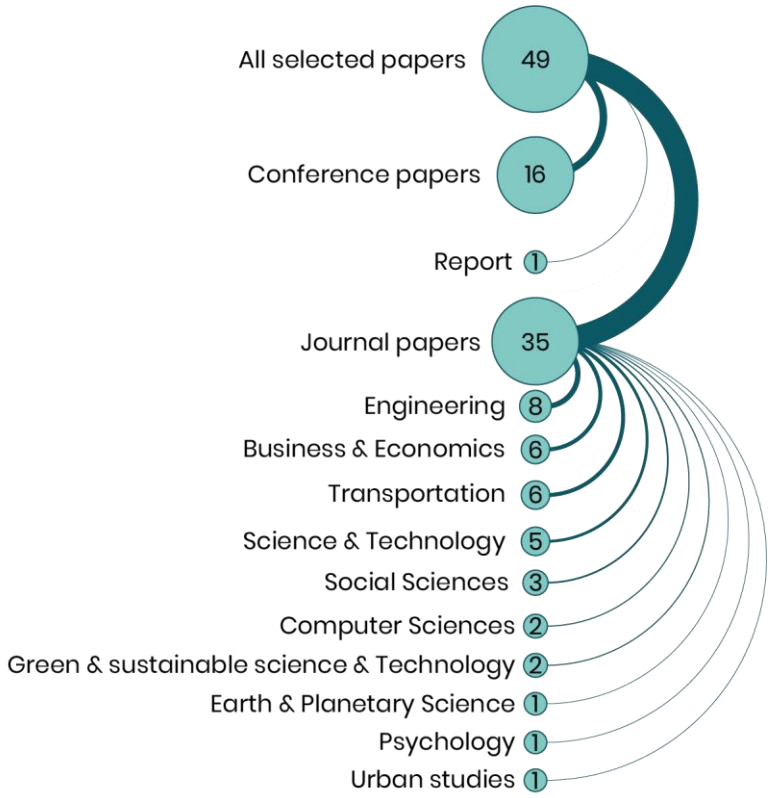


Figure 6. Types of publication and journal fields.

Concerning the geographical distribution of publications, the total number of authors is 144 (*Table 2*). Out of the total, 41% are affiliated with European institutions, 37% with Asian institutions, 22% with American institutions, and 1% with Oceanian institutions. Europe has the largest coverage among the continents. In Europe, Italy constitutes 44% of the selected publications, followed by Poland (12%), Portugal (10%), the United Kingdom (8%), Spain and the Netherlands (7%), Croatia (3%), Austria, Belgium, Romania, France, and Switzerland (2%). In Asia, China constitutes 42% of the selected publications, followed by India (17%), Malaysia and Bangladesh (9%), Taiwan (8%), Iran (6%), Turkey and South Korea (4%), and Iraq (2%). In America, Brazil constitutes 52% of the selected publications, followed by Colombia (32%), the USA (10%), and Chile (6%). In Oceania, an author is affiliated with a New Zealand institution. *Table 3* shows the geographic distribution by continents, countries of the case studies treated, and the focus on the different categories and modes of transport. The analysis reports that 67 case studies were distributed as follows: 52% in Europe, 27% in Asia, 19% in America, and 2% in Oceania. Finally, one article does not present a specific case study (2%). 55% of the case studies concern non-motorized mobility, 34% multimodality, 9% public transport, and 2% electric micro-mobility. A reasonable part of the case studies investigated the infrastructural aspects, especially concerning pedestrian mobility. It is interesting to note that in America and Asia, there is a prevailing interest in non-motorized mobility. In Europe, interest is also largely in LPT, as well as pedestrian and cycling mobility. In terms of multimodality, most of the studies

included non-motorized mobility and LPT services, which were investigated in Europe and Oceania.

Table 2. Geographical distribution of the publications reviewed.

Continent	N° of publications	N° of researchers involved	Researchers' country	N° of researchers involved
America	10	31	Brazil	16
			Chile	2
			Colombia	10
			USA	3
Asia	17	53	Bangladesh	5
			China	22
			India	9
			Iran	3
			Iraq	1
			Malaysia	5
			South Korea	2
			Taiwan	4
			Turkey	2
Europe	21	59	Austria	1
			Belgium	1
			Croatia	2
			France	1
			Italy	26
			Netherlands	4
			Poland	7
			Portugal	6
			Romania	1
			Spain	4
			Switzerland	1
			UK	5
Oceania	1	1	New Zealand	1
Total	**49**	**144**	**Total**	**144**

Table 3. Geographic distribution of case study analysed in the publications, divided by transport category.

Geographical area	Local Public Transport	Electric micro-mobility	Non-motorized mobility	Multimodal transport	Total
America	**1**	**0**	**11**	**0**	**12**
Brazil	1	0	3	0	4
Colombia	0	0	4	0	4
USA	0	0	4	0	4
Asia	**0**	**0**	**15**	**2**	**17**
Bangladesh	0	0	1	0	1
China	0	0	5	0	5
India	0	0	3	2	5
Iran	0	0	1	0	1
Iraq	0	0	1	0	1
Malaysia	0	0	1	0	1
Taiwan	0	0	2	0	2
Turkey	0	0	1	0	1
Europe	**5**	**1**	**8**	**19**	**33**
Croatia	1	0	0	0	1
Greece	0	0	0	2	2
Ireland	0	0	0	2	2
Italy	2	0	4	3	9
Lithuania	0	0	0	2	2
Netherlands	0	0	0	1	1
Poland	1	1	1	0	3
Portugal	1	0	1	0	2
Romania	0	0	0	2	2
Spain	0	0	1	3	4
Sweden	0	0	0	2	2
UK	0	0	1	2	3
Oceania	**0**	**0**	**0**	**1**	**1**
New Zealand	0	0	0	1	1
Not present	**0**	**0**	**1**	**0**	**1**
Total	**6**	**1**	**35**	**22**	**64**

As illustrated in *Figure 7*, the quality of non-motorized mobility has consistently attracted scholarly attention, with a notable surge in publications in 2019 and 2020. This increase is likely linked to the COVID-19 pandemic, which highlighted the importance of individual and personally distanced modes of transport. Conversely, micro-electric mobility, which is a relatively recent development, remains underrepresented in the literature, reflecting its more recent emergence in urban mobility systems.

Interest in multimodal mobility and the integrated assessment of transport systems has grown steadily over the past decade, although the overall volume of research has not expanded at the same rate. *Figure 7*, also shows that studies on non-motorized mobility appeared earlier than those focusing on LPT. However, this finding primarily results from the deliberate exclusion of service-quality studies, which, as previously discussed, were among the earliest and most extensively investigated within the transport quality domain.

In line with this, *Figure 8* traces the evolution of research on LPT in its various forms, confirming that LPT services have received the most sustained scholarly attention, with some seminal contributions dating back to the early 2000s and earlier.



Figure 7. Number of publications reviewed per year and transport categories.

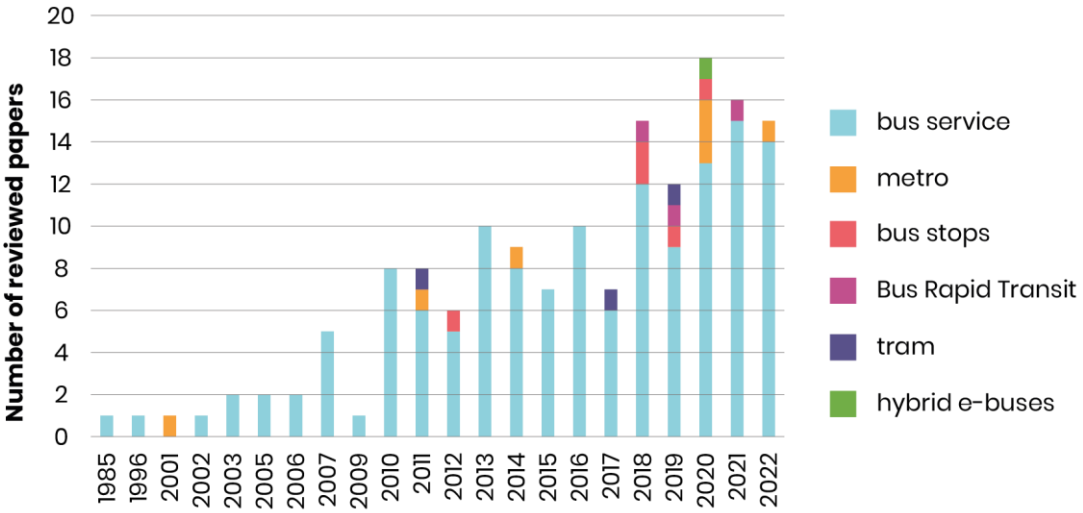


Figure 8. Number of publications reviewed on LPT per year and public transport typology.

2.2.2 Evolution of the quality concept in transport systems

Transport quality is often described using various terms, such as “performance measures”, “Level of Service” (LOS), and “Quality of Service”, which are frequently used interchangeably. Nevertheless, these terms denote distinct concepts, as clarified in the Transit Capacity and Quality of Service Manual (TCQSM) (Kittleson & Associates, 1999).

Performance measures in transport planning emerged during the mid-20th century as a response to the rapid expansion of the transport sector. These measures focus on quantifiable, objective metrics such as traffic flow, travel time, vehicle count, capacity, and speed to evaluate the efficiency and effectiveness of transport infrastructure (Meyer, 1995). The primary goal was to assess system performance and optimize infrastructure to accommodate increasing volumes of vehicles and traffic. This approach laid the foundation for subsequent developments in transport planning, which later evolved into the LOS and QoS frameworks.

The concept of LOS was first introduced in the 1950s, primarily focusing on vehicular traffic and using metrics like vehicle density and speed to assess road performance (Highway Research Board, 1965). It is defined as a classification system that ranges from “A” (best) to “F” (worst), representing six designated values for a specific aspect of service or infrastructure. Engineering-based perspectives heavily influenced this approach, aiming to optimize traffic flow.

In the late 1970s and early 1980s, the limitations of the LOS framework began to draw critical attention, particularly regarding its neglect of user perceptions and experiences. In response, LOS-type evaluations have been tested by shifting the focus to the user's perspective (Botzow, 1974; Pattnaik & Ramesh Kumar, 1996).

During the 1990s, there was a shift from the LOS to the Quality of Service (QoS) concept, which emphasized passenger-perceived quality and included factors such as comfort,

safety, information, and accessibility (Pullen, 1993). The introduction of QoS frameworks, especially through institutional tools like the TCQSM, marked a pivotal shift from traditional infrastructure performance to user-centred service quality assessment. This transition from LOS to QoS represents a broader reorientation in transport planning, shifting from purely technical evaluations to more holistic, user-experience-driven assessments (de Oña & de Oña, 2014). While LOS emphasizes objective operational metrics such as speed and capacity, QoS frameworks incorporate subjective factors, enabling planners to evaluate how services are experienced by those who use them.

With the growing emphasis on sustainable transport, these evaluation frameworks are increasingly being extended to encompass non-motorized mobility, giving rise to specialized indices such as the Pedestrian Level of Service (PLOS) (Fruin, 1971) and the Quality of Pedestrian Level of Service (QPLOS) (Talavera-Garcia & Soria-Lara, 2015) for pedestrians, and the Bicycle Level of Service (BLOS) for cyclists (Transportation Research Board, 1997), which, over time, have also been related to land use. Recent studies have further expanded the application of PLOS beyond traditional pedestrian environments to intermodal settings such as metro stations. Hu et al. (2019) propose an analytical framework that evaluates pedestrian performance across key functional areas within transit stations, introducing the Interaction Index (II) to assess both efficiency and safety. This approach underscores the relevance of pedestrian infrastructure not only in isolation but also as a fundamental component of multimodal transport systems. Ujjwal and Bandyopadhyaya (2021) integrated both qualitative and quantitative parameters in the evaluation of sidewalk conditions in congested, mixed land-use locations, identifying relevant assessment factors for urban areas with such characteristics. Building upon this foundation, their later work (2022) introduced a comprehensive service quality assessment framework that explicitly considers the gap between pedestrians' desired and perceived walking conditions. This model expands previous approaches by incorporating a wide range of built and walking environment-related variables and uses the Fuzzy Hierarchy Process to determine the relative importance of each factor. By including both objective metrics and user expectations, the framework offers a more refined and user-centred method for evaluating pedestrian infrastructure.

Until then, few studies had integrated these aspects into PLOS assessments. As sustainable and multimodal transportation became central to urban mobility strategies, the Multimodal Level of Service (MMLOS) framework was introduced by AASHTO and the FHWA in 2008, integrating pedestrian and bicycle LOS within a unified methodology (NASEM, 2008). However, MMLOS remains limited in scope, excluding micromobility vehicles and environmental quality aspects outside the right-of-way. More recently, with the spread of electric and shared micromobility systems, quality assessment has also been applied to more innovative and technology-based systems, such as Scooter Level of Service (SLOS) for e-scooters (Kazemzadeh and Sprei, 2022). In shared micromobility, Hamerska et al. (2022) proposed the MicroMobility QUALity (MMQUAL) model to evaluate

the quality of e-scooter systems based on application design, hardware features, and customer support.

This evolution reflects a growing recognition of the importance of both technical and user-centred measures of service quality. For example, Eboli and Mazzulla (2011) proposed an integrated methodology for evaluating LPT service quality, combining both objective indicators (e.g., service frequency, travel time) and subjective perceptions collected from passengers. Their approach emphasizes the need to understand service quality not only through technical performance but also through user experience. The results showed that while objective factors such as punctuality and frequency significantly influence user satisfaction, subjective measures are crucial for capturing the overall perception of service quality. This dual-perspective analysis allows better target service improvements based on what matters most to passengers. Rodriguez-Valencia et al. (2020) examined pedestrian QoS on urban sidewalks, revealing that user perceptions significantly shape how service quality is experienced, highlighting the value of subjective assessments alongside objective indicators. Similarly, Ashqar et al. (2022) applied a QoS framework to bike-sharing systems by introducing the concept of Optimal Occupancy, which enabled more accurate predictions of service performance and more efficient station placement through spatial modelling.

These newer approaches resonate with the earlier SERVQUAL model proposed by Parasuraman et al. (1988), who emphasized the role of user expectations and perceptions in evaluating service quality across five dimensions: reliability, assurance, tangibles, empathy, and responsiveness. Although initially developed for general service industries, SERVQUAL later gained traction in the transport field as scholars sought tools to assess transit users' subjective experiences better. This evolution enabled more accurate measurement of perceived transport service quality. Studies by de Oña & de Oña (2014) applied SERVQUAL dimensions to metro systems in European cities using advanced methods such as Structural Equation Modelling (SEM) to analyse the determinants of satisfaction and behavioural intentions. Bakti and Sumaedi (2015) developed the P-TRANSQUAL model, tailoring the original dimensions to factors more directly aligned with the experiences of LPT users, such as comfort, tangibility, personnel, and reliability.

In line with these subjective approaches, Redman et al. (2012) stress that PTCs can more effectively encourage modal shifts from private cars by focusing on the quality attributes perceived as valuable by car users. Instead of implementing technical improvements without user context, PTCs should identify the motivations behind car use - e.g., flexibility, emotional attachment, or perceived convenience - and design services that replicate or offset these benefits through enhanced user experiences. This implies prioritizing not only basic functional attributes like reliability and accessibility but also context-specific, affective elements, such as station aesthetics, perceived safety, and quality of information.

Building on this user-centred perspective, evaluation models that account for both subjective and objective dimensions can serve as powerful tools to guide practical

improvements. As shown by Marisamynathan and Lakshmi (2016), such models can be applied to a wide range of use cases, including the assessment of existing infrastructural conditions, the enhancement of users' safety and comfort, and the optimization of infrastructure geometry in response to the surrounding road environment.

Moreover, habit disruption strategies - e.g., limited-time free tickets or targeted promotional events - can be effective in prompting car users to try LPT, provided the service quality is perceived as competitive or superior. Sustaining this shift depends on whether the system continuously delivers perceived value on the attributes users consider most critical. This approach reinforces the growing consensus that quality assessment frameworks must reflect user-specific values and perceptions, especially when aiming to shift deeply ingrained car-oriented behaviours.

A step forward in this type of assessment was made by Barabino and Di Francesco (2016), who introduced a holistic framework called TRANSQUAL for evaluating the quality of LPT service. The key innovation is the use of percentage-based metrics to assess service quality from the passenger's perspective. Specifically, this method allows for measuring the percentage of passengers: (i) who expect a predefined level of service; (ii) for whom the service is designed; (iii) who receive the planned service; (iv) who perceive the service as delivered. Collectively, these metrics are used to identify gaps and recommend corrective actions, emphasizing a passenger-centred approach to assessing service quality.

The evolution of the concept of quality in LPT has resulted in the emergence of multiple approaches and perspectives, including those of PTCs, shared mobility service providers (e.g., bike-sharing and scooter-sharing companies), the community, and passengers (Kittelson & Associates, 2013). From the perspective of PTCs, transit quality is primarily evaluated as a business asset. In this context, the European standard EN 13816:2002 introduces an innovative framework for the measurement and management of QoS, enabling PTCs to subject their routes to external certification (CEN/TC 320, 2002). The community perspective, particularly relevant to Public Administrations, assesses transit systems in terms of their contributions to broader societal objectives, such as improving overall mobility and reducing environmental impact. Finally, the passenger perspective centres on the overall measured or perceived performance of transit service from the passenger's viewpoint (Transportation Research Board, 2003). *Figure 9* provides a summary of the main quality assessment methods in the transportation field.

Yet, despite this broadening scope, there remains no universally accepted definition of transport quality for sustainable and multimodal systems. Existing definitions are mostly sectoral and refer mainly to the service quality of LPT. The broadest and most general definition describes quality as the most synthetic, essential, and comprehensive attribute of transport system performance (Raicu & Raicu, 2005). Barabino & Di Francesco (2016) define service quality as the capability of transit operators to provide a service that reflects current and potential user desires. Liou et al. (2014) and Hsu et al. (2018) describe it as a

key factor influencing people's willingness to use public rather than private transport. de Oña and de Oña (2014), and Julio et al. (2022) regard it as an essential tool for public transport agencies and planners to attract and retain passengers.

Several authors link the quality of the urban environment with the pedestrian transport system. For instance, Talavera-García & Soria-Lara (2015) introduce the Q-PLOS and associate quality with the urban design and its relationship to walking needs. While they do not explicitly define the quality of transport systems, they relate the quality of pedestrian environments to inhabitants' satisfaction with walking conditions. Sousa et al. (2019) argue that walkable environments contribute to environmental quality, applying the Pedestrian Environment Quality Index (PEQI) to assess medium-sized European cities. Diana et al. (2016) emphasize that the measurement of quality should reflect the entire travel experience, not merely the use of a specific service.

Main sources	Index/Method	Transport modes	Quality aspects		Evaluation		Perspective		
			Se	In	Ob	Su	Us	PTC	Go
HRB, 1965	LOS	Road traffic		●	●			●	●
Botzow, 1974	LOS	LPT	●		●		●		
Pattnaik & Kumar, 1996	LOS	LPT	●			●	●		
Pullen, 1993	LOS	LPT	●		●	●	●		
CEN/TC, 2002	QoS	LPT	●		●	●		●	
TRB, 1999; 2003	QoS	LPT	●			●	●		
de Oña & de Oña, 2014	QoS	LPT	●			●	●		
Ashqar et al., 2022	QoS	Bike-sharing	●			●	●		
Fruin, 1971	PLOS	Pedestrian	●	●	●	●	●		
Rodriguez-Valencia et al., 2020	PLOS	Pedestrian	●	●		●	●		
Talavera-García & Soria-Lara, 2015	Q-PLOS	Pedestrian	●	●	●	●	●		
Ujjwal & Bandyopadhyaya, 2021	PLOS	Pedestrian	●	●	●	●	●		
TRB, 1997	BLOS	Cycling mobility	●	●	●		●		
NASEM, 2008	MMLOS	Cars, LPT, Ped., Cyc.	●	●	●			●	●
Kazemzadeh & Sprei, 2022	SLOS	e-Micromobility	●	●	●		●		
Hamerska et al., 2022	MMQUAL	e-Micromobility	●			●	●		
Parasuraman et al., 1988	SERVQUAL	Multi-sectors	●			●	●		
de Oña & de Oña, 2014	SERVQUAL	Metro	●			●	●		
Bakati & Sumaedi, 2015	P-TRANSQUAL	LPT	●			●	●		
Barabino & Di Francesco, 2016	TRANSQUAL	LPT	●		●	●	●	●	

Se service **Ob** objective **Us** users

In infrastructure **Su** subjective **PTC** Public Transport Company

Go Government

Figure 9. A summary of the approaches to measure Quality in Transportation. The papers are listed in chronological order, from the earliest to the most recent.

2.2.3 Rationale and purpose of transport quality assessment

This section explores the importance of measuring the quality of sustainable transport systems and identifying potential future directions for research in this area.

Several pieces of evidence emerged based on the highlighted transport system. The following sub-sections highlight the main quality-related aspects of the transport systems investigated in this paper and summarize common features that characterize the quality of multimodality. These aspects include the relevance of quality assessment, dimensions of quality, implications for different stakeholders, research gaps, and future directions.

Pedestrian mobility

Assessing the quality of pedestrian mobility is essential to capture the perceptions of pedestrians using pedestrian transportation infrastructure. To achieve this, it is necessary to consider both tangible and intangible factors including sidewalk infrastructure, safety, environmental context, and social interactions (Vallejo-Borda et al., 2020; Jahan et al., 2020; Soltani et al., 2017).

Quality assessment allows for a more comprehensive understanding of the full complexity of walking experiences by combining objective indicators (e.g., slope, surface quality, sidewalk width) with subjective assessments and stated preference techniques (Sousa et al., 2019; Bellizzi et al., 2019; Da Rocha et al., 2019). User perception plays a crucial role in determining walkability, sometimes outweighing operational measures like sidewalk width or traffic volume (Rodriguez-Valencia et al., 2020).

Factors such as safety and security (Ujjwall & Bandyopadhyaya, 2021), comfort (Sarker, 2003), and aesthetic and environmental quality (Bellizzi et al., 2021; Blečić et al., 2016) emerge as critical dimensions in the evaluation of pedestrian infrastructure, with some models even integrating psychosocial variables such as fear of crime and emotional discomfort (Arellana et al., 2019; Vallejo-Borda et al., 2020).

There is growing interest in using latent variable modelling, cognitive mapping, and advanced logit models to better interpret pedestrian preferences and the heterogeneity of walking behaviour (Vallejo-Borda et al., 2020; Bellizzi et al., 2021; Larranaga et al., 2018).

Future research directions should include: validating existing QoS models across different urban contexts and population groups (e.g., children, elderly, women); exploring the temporal dynamics of pedestrian perceptions (e.g., day vs. night conditions); integrating emotional states and individual characteristics (Vallejo-Borda et al., 2020; Ismail et al., 2020); and developing interactive tools such as walkability maps or GIS-integrated audits to support planning and policy-making (Sarker, 2003; Arellana et al., 2019). Additionally, replicating methodologies in Global South contexts (Larranaga et al., 2018; Da Rocha et al., 2019) and incorporating emerging micromobility interactions (e.g., bike-pedestrian

conflicts) remains a critical challenge for improving the inclusivity and resilience of pedestrian environments in future cities.

Figure 10 provides a summary of the main evidence from the quality assessment of pedestrian mobility that emerged from the literature review.

Dimentions	Key points	Implications for practice	Main sources
Relevance of quality assessment	Quality assessment captures the full complexity of walking experiences, revealing how perceptions often predict walkability better than objective measures.	Quality assessment helps tailor infrastructure to users' real needs and promotes more human-centred urban design.	Vallejo-Borda et al., 2020. Jahan et al., 2020. Soltani et al., 2017. Rodriguez-Valencia et al., 2020.
Quality dimensions identified	Sidewalk design, safety, comfort, accessibility, aesthetics, and environmental/social context as core quality dimensions.	Enable comprehensive evaluation of pedestrian areas and prioritization of improvement actions.	Ujjwall & Bandyopadhyaya, 2021. Sarker, 2003. Bellizzi et al., 2021. Blečić et al., 2016. Arrellana et al., 2019
Stakeholder implications	Encourage planners and policymakers to consider user perceptions (e.g., fear of crime, comfort, environmental quality) alongside technical standards to design more inclusive public spaces.	Apply perception-based tools (e.g., surveys, walkability audits). Strengthen stakeholder engagement to align design choices with community needs.	Da Rocha et al., 2019. Ismail et al., 2020. Rodriguez-Valencia et al., 2020.
Research gaps & future directions	Validate existing models across contexts and user groups. Integrate emotional & temporal factors. Develop advanced modelling and GIS-based tools.	Support evidence-based planning. Promote equity-oriented infrastructure design. Enhance methodological robustness through context-sensitive approaches.	Vallejo-Borda et al., 2020. Larranaga et al., 2019. Arrellana et al., 2019. Bellizzi et al., 2021.

Figure 10. Summary of the main evidence from the assessment of the quality of pedestrian mobility.

Cycling mobility

Assessing the quality of cycling mobility is essential to support the shift toward more sustainable and inclusive urban mobility systems. The literature reveals that infrastructure quality affects both actual usability and perceived comfort, which are critical for long-term adoption. For instance, Hsu & Lin (2011) demonstrate that many existing street segments require significant redesign to become bike-appropriate, highlighting the gap between infrastructure availability and suitability. Similarly, Calvey et al. (2015) show that surface conditions and maintenance play a decisive role in shaping cyclists' comfort and satisfaction. Looking ahead, future research should move beyond isolated quality indicators and develop transferable, context-aware assessment frameworks. As suggested by de Aquino Traldi et al. (2022), incorporating spatial interpolation and multi-

criteria weighting methods allows for broader network-wide evaluations while remaining sensitive to local variations. Additionally, Fistola et al. (2020) emphasize the importance of integrating cycling infrastructure assessment with multimodal planning, particularly in terms of connectivity with LPT and high-demand urban areas. Research should also invest in validating assessment tools across diverse urban contexts and user profiles and explore real-time monitoring techniques that can dynamically inform infrastructure maintenance and upgrade priorities. Ultimately, quality assessment should not only guide technical improvements but also serve as a strategic planning tool to create safer, more accessible, and better-integrated cycling environments.

Figure 11 provides a summary of the main evidence from the quality assessment of cycling mobility that emerged from the literature review.

Dimentions	Key points	Implications for practice	Main sources
Relevance of quality assessment	Quality assessment is essential to evaluate the effectiveness of cycling infrastructure, supporting low-emission and space-efficient urban mobility. It helps identify gaps between perceived and actual network suitability, guiding infrastructure design and improvement.	Quality assessment enables targeted interventions and supports safe, sustainable cycling adoption.	Hsu & Lin, 2011. Calvey et al., 2015. Fistola et al., 2020.
Quality dimensions identified	Geometric design, pavement quality, connectivity, comfort, and integration with public transport define cycling quality.	Support multi-criteria evaluation and holistic planning of cycling networks.	Hsu & Lin, 2011. Calvey et al., 2015. Traldi et al., 2022. Fistola et al., 2020.
Stakeholder implications	Support planners, municipalities, and operators in integrating real-time monitoring, spatial analysis, and user feedback into design and policy decisions.	Promote participatory planning and coordination between cycling, land use, and public transport policies. Foster user-centred design and improved multimodal accessibility.	Calvey et al., 2015. Fistola et al., 2020.
Research gaps & future directions	Refine quality models for local conditions and users. Integrate cycling with LPT and land-use systems. Expand geo-spatial evaluations.	Generalize models across contexts and refine performance indicators. Link infrastructure quality to actual usage data. Strengthen multimodal planning.	Traldi et al., 2022. Fistola et al., 2020.

Figure 11. Summary of the main evidence from the assessment of the quality of cycling mobility.

Bike-sharing

The increasing integration of bike-sharing systems (BSS) into urban mobility networks has intensified the need for robust methods to assess their service quality. Assessing the quality of BSS is essential not only for identifying the factors that influence user satisfaction and system usage but also for informing urban transport planning and policy aimed at promoting sustainable and multimodal mobility. High-quality bike-sharing services can contribute to reducing car dependency, improving urban air quality, and enhancing first- and last-mile connectivity within urban mobility systems.

The literature highlights a broad set of quality criteria that users associate with satisfactory BSS, including convenience, accessibility, comfort, safety, affordability, infrastructure quality, and system reliability (Xue et al., 2022; Liang et al., 2021; Julio et al., 2022). Several studies emphasize that perceived quality is highly context-dependent and varies with geographic, climatic, and socio-cultural conditions, as well as temporal factors such as weekdays versus weekends or peak versus off-peak hours (Xue et al., 2022; Shin, 2020). Additionally, service quality perceptions differ across stakeholder groups; in fact, Public Administrations (PAs), Public Transport Companies (PTCs) service operators, and users often have perspectives, which may lead to ineffective policy or operational decisions (Ma et al., 2019). Since users are the ones who use a transportation service or infrastructure, it is their perception that should take priority if we want to support evidence-based improvements and promote trust in the service.

Another key rationale for quality evaluation is its strategic value in optimising systems and allocating resources. By identifying which service attributes, such as bicycle availability, pricing clarity, technical reliability of stations, or responsiveness of customer service, are most influential in shaping user satisfaction, operators, policy-makers and planners can prioritize interventions that are both cost-effective and impactful (Julio et al., 2022; Shin, 2020).

From a research perspective, future studies are encouraged to move beyond static and operator-driven assessments. First, there is a need for more nuanced and scalable evaluation frameworks, including multi-criteria decision-making (MCDM) models, fuzzy logic, and Grey Theory approaches, which allow for the integration of subjective user preferences and uncertainty (Hsu et al., 2018; Liang et al., 2021). These methods can provide more comprehensive insights compared to traditional statistical techniques. Second, researchers should incorporate emerging data sources, such as social media content, into quality evaluations to capture real-time feedback and shifting public sentiment (Shin, 2020).

Further research should also address the heterogeneity of users, exploring how demographic characteristics (e.g., gender, age, income level) affect quality perceptions (Ma et al., 2019); it also should consider the role of public authorities in shaping citizens' awareness, knowledge, and acceptance of BSS (Podgorniak-Krzykacz et al., 2022).

Longitudinal studies are particularly needed to monitor how service quality evolves in response to policy changes or technological innovations.

In summary, BSS quality evaluation is not merely a technical exercise but a critical component of sustainable transport planning. It enables more inclusive and user-responsive system design and plays a foundational role in the integration of bike-sharing within multimodal urban mobility strategies.

Figure 12 provides a summary of the main evidence from the quality assessment of bike-sharing systems (BSS) that emerged from the literature review.

Dimentions	Key points	Implications for practice	Main sources
Relevance of quality assessment	Quality assessment enhances user satisfaction, exposes service gaps and stakeholder misalignments, and informs multimodal planning.	Quality assessment supports user-driven planning and adaptation of services to local contexts and diverse needs.	Xue et al., 2022. Shin, 2020. Ma et al., 2019.
Quality dimensions identified	Accessibility, availability, safety, reliability, pricing, station performance, and customer support determine service quality.	Guide investment toward features with highest impact on perceived quality and usage.	Liang et al., 2021. Julio et al., 2022. Shin, 2020.
Stakeholder implications	Highlight perception gaps between users, operators, and public authorities, emphasizing the need for participatory quality assessment.	Engage users and local stakeholders in co-design and evaluation to align expectations and improve service quality.	Ma et al., 2019. Podgorniak-Krzykacz et al., 2022.
Research gaps & future directions	Apply MCA and fuzzy approaches for richer evaluation. Incorporate social media and user-generated data sources. Explore cross-city comparisons.	Develop adaptive and data-driven planning tools. Ensure long-term relevance of evaluations across user groups. Promote responsive service design.	Hsu et al., 2018. Shin, 2020. Ma et al., 2019.

Figure 12. Summary of the main evidence from the assessment of the quality of bike-sharing systems (BSS).

Electric micromobility

Assessing the quality of electric micromobility is essential to ensuring that these emerging transport options realize their full potential as sustainable first- and last-mile solutions in urban mobility systems. As shared electric scooters become increasingly popular, they offer alternatives to car use and support goals related to congestion reduction and environmental sustainability. However, the effectiveness of electric micromobility depends not only on its availability, but also on the quality of the user experience. Without reliable, accessible, and intuitive services, these systems risk being underutilized or

generating negative externalities such as clutter, unsafe driving practices, or unequal access (Hamerska et al., 2022).

Currently, the lack of dedicated and specific quality assessment frameworks for shared electric scooters is a significant research gap. Filling this gap is essential to understanding which aspects of the service most influence user satisfaction and long-term adoption. The MMQUAL (MicroMobility QUALity) model proposed by Hamerska et al. (2022) provides a theoretical first step by identifying three key dimensions: mobile application functions, device characteristics, and customer service. These dimensions determine perceived quality and indicate how in digitally oriented mobility services, app functionality is central to user experience.

Looking ahead, future research should focus on validating and refining these qualitative dimensions in different urban and cultural contexts, ensuring the robustness and generalizability of the model. In addition, additional dimensions such as environmental impact, urban integration, and regulatory compliance may need to be incorporated. The development of dynamic, user-cantered, and scalable assessment tools will be critical to guide evidence-based policies and improve system design. Ultimately, quality assessment in electric micromobility is not just a technical necessity, but a strategic tool for aligning operator performance, user expectations, and public interest in the transition to more sustainable urban transport.

Figure 13 provides a summary of the main evidence from the quality assessment of electric micromobility that emerged from the literature review.

Dimentions	Key points	Implications for practice	Main sources
Relevance of quality assessment	Quality assessment ensures electric micromobility effectively serve first/last-mile needs, highlights the lack of tailored evaluation frameworks, and supports the development of models tailored to the specific features of e-scooter services.	Quality assessment highlights the need for specific quality models tailored to shared e-scooter systems.	Hamerska et al., 2022.
Quality dimensions identified	App usability, device characteristics, and customer service form the main quality dimensions of shared e-scooter systems.	Provide a structured framework (e.g., MMQUAL) for assessing and enhancing e-micromobility services.	Hamerska et al., 2022.
Stakeholder implications	Enable operators to benchmark and enhance performance using structured models such as MMQUAL, focusing on both app usability and customer support.	Direct strategic investments toward digital interfaces and vehicle quality to strengthen user satisfaction and competitiveness.	Hamerska et al., 2022.
Research gaps & future directions	Validate and adapt quality models (e.g., MMQUAL) in different contexts. Extend assessment to environmental and regulatory aspects, establishing frameworks for continuous improvement.	Build foundations for broader, scalable evaluation models. Prioritize model validation and long-term refinement, supporting strategic development of emerging mobility services.	Hamerska et al., 2022.

Figure 13. ummary of the main evidence from the assessment of the quality of electric micromobility.

Local Public Transport

Assessing the quality of LPT bus stops is essential to improve a sustainable and multimodal urban mobility. The reviewed literature emphasizes that passengers' satisfaction and their willingness to use LPT are strongly influenced by a combination of physical, functional, and psychological attributes of the waiting environment. Studies such as Ely et al. (2012) highlight the importance of thermal comfort, weather protection, and seating, suggesting that seasonal variations and demographic differences (e.g., tourists) should be considered in future evaluations. Other contributions (e.g., Barabino, 2018; Rossetti, 2020) have moved towards operational frameworks capable of identifying performance deficiencies at the level of individual stops, often using in-field observations and spatial analysis tools. The need to adapt infrastructure to specific user groups, such as elderly or disabled passengers, emerges clearly in studies like Rosa & Da Cruz Lopes (2019) and Podciborski (2017), underlining accessibility, safety, and clarity of information as essential components of inclusive design. Furthermore, while traditional service quality models (e.g., Vujičić & Prester, 2019) provide valuable performance benchmarks, they

often lack integration with environmental or contextual data, which future research should increasingly consider. Overall, these results highlight the need to develop multidimensional, user-centred assessment tools that combine subjective perception with objective data. Future studies should aim to generalize existing methodologies to apply to all cities and transport modes, incorporate dynamic and seasonal factors, and expand the role of inclusive design approaches in quality assessment.

Figure 14 provides a summary of the main evidence from the quality assessment of LPT bus stops that emerged from the literature review.

Dimentions	Key points	Implications for practice	Main sources
Relevance of quality assessment	Quality assessment is essential to evaluate how stop infrastructure influences passenger satisfaction, comfort, and modal choice. It supports service improvement, guides resource allocation, and strengthens sustainable and multimodal mobility strategies.	Quality assessment guides resource allocation and supports structured, user-focused quality assessments.	Ely et al., 2012. Podciborski, 2017. Rossetti, 2020.
Quality dimensions identified	Shelter, comfort, accessibility, safety, information, and context-sensitive design define stop and station quality.	Offer a multidimensional checklist for LPT stop design and operational monitoring.	Barabino, 2018. Rosa & Da Cruz Lopes, 2019. Vujičić & Prester, 2019.
Stakeholder implications	Assist local authorities and transit operators in prioritizing interventions and ensuring inclusive design for elderly, disabled, and tourist users.	Encourage participatory and universal design approaches to align service quality with diverse user needs.	Rosa & Da Cruz Lopes, 2019. Rossetti, 2020. Podciborski, 2017.
Research gaps & future directions	Combine subjective and objective indicators for a comprehensive assessment, as well as different dimensions (environmental, seasonal,..). Generalize frameworks across cities and transport modes.	Adopt adaptable, multimodal, and user-centred tools. Apply dynamic evaluation methods responsive to context.	Barabino, 2018. Vujičić & Prester, 2019. Ely et al., 2012.

Figure 14. Summary of the main evidence from the assessment of the quality of LPT infrastructure.

Multimodal quality

Assessing the quality of multimodal transport systems is essential to develop integrated, efficient, and user-centric urban mobility networks. A key reason lies in the complexity of the user journey, which increasingly involves transitions across multiple transport modes: walking, bike-sharing, e-micromobility, and LPT. These transitions often occur at multimodal transport hubs (MMTHs), where the quality of the transfer environment, accessibility, safety, and comfort significantly influences overall satisfaction and mode choice (Chauhan et al., 2021; Groenendijk et al., 2018). A poor-quality interface can discourage users from using sustainable alternatives, whereas seamless multimodal integration can reduce reliance on private vehicles and support inclusive mobility (Talavera-Garcia & Soria-Lara, 2015).

Moreover, multimodal quality assessment provides a framework to detect barriers - both physical and perceptual - that limit the effectiveness of transport networks. Tools like Q-PLOS (Talavera-Garcia & Soria-Lara, 2015), GIS-based evaluations (D'Orso & Migliore, 2019), and the Node-Place-Experience (NPE) model (Groenendijk et al., 2018) highlight the importance of combining objective indicators with subjective perceptions such as environmental pleasantness. These tools help transport planners to prioritize interventions that enhance first- and last-mile connectivity, walkability near hubs, and the overall traveller experience.

In addition, studies show that local context and user diversity strongly influence perceived quality. Differences in needs and expectations across neighbourhoods, social groups, and cultural settings require quality assessments to be adaptable and inclusive (Burlando et al., 2021; Diana et al., 2017). Therefore, multimodal quality evaluation is not just a technical measurement but a strategic tool to guide context-sensitive and user-responsive planning (Douglas, 2015; Chauhan et al., 2021).

Figure 15 provides a summary of the main evidence from the quality assessment of multimodal transport that emerged from the literature review.

Dimentions	Key points	Implications for practice	Main sources
Relevance of quality assessment	Quality assessment is essential to understand the complexity of user journeys involving multiple transport modes. It supports the design of integrated, efficient, and user-centred mobility systems, encouraging sustainable travel choices.	Quality assessment uses user feedback to enhance multimodal design and operational coordination.	Chauhan et al., 2021. Groenendijk et al., 2018. Douglas, 2015.
Quality dimensions identified	Accessibility, safety, signage, comfort, transfer environment, and environmental integration shape multimodal quality.	Combine objective and perceived dimensions to design inclusive and seamless multimodal systems.	Chauhan et al., 2021. Diana et al., 2016. Burlando et al., 2021. Groenendijk et al., 2018. Talavera-Garcia & Soria-Lara, 2015.
Stakeholder implications	Guide public administrations, operators, and urban designers in coordinating actions around user-centred, context-sensitive multimodal networks.	Support user-based planning and interventions tailored to different contexts and demographic groups.	Diana et al., 2017. Burlando et al., 2021. D'Orso & Migliore, 2019. Talavera-Garcia & Soria-Lara, 2015.
Research gaps & future directions	Validate multimodal quality frameworks across cities and cultures. Integrate objective performance with subjective perception, including social and psychological aspects in evaluation models.	Expand empirical testing of multimodal frameworks. Build scalable and flexible tools for different urban contexts, aligning research with inclusive and cross-modal planning needs.	Chauhan et al., 2021. Groenendijk et al., 2018. Diana et al., 2016. Burlando et al., 2021.

Figure 15. Summary of the main evidence from the assessment of the quality of multimodal transport.

Evaluating quality in multimodal transport systems is crucial for enhancing the integration across different transport modes, incentivizing users to use more sustainable transport systems than the private car. The literature review highlights that each transportation system has its specific dimensions of quality. However, they all converge around key themes such as accessibility, safety, comfort, system reliability and user perception (Xue et al., 2022; Bellizzi et al., 2021; Liang et al., 2021; Podciborski, 2017). In multimodal contexts, these elements become even more significant because transitions between transport modes can determine the overall success and adoption of sustainable transport solutions. Some examples of these transitions could be the shift from bike-sharing to LPT or from electric scooters to walking (Fistola et al., 2020; Julio et al., 2022; Vallejo-Borda et al., 2020).

Quality assessment enables the identification of physical and psychological barriers in modal interfaces, enabling transport planners to reduce critical issues and improve the overall travel experience. For example, studies on LPT and bike-sharing underline the strategic role of first- and last-mile connectivity (Fistola et al., 2020; Shin, 2020), while e-

micromobility research points to the usability of digital platforms and mobile apps as part of the user journey (Ma et al., 2019; Liang et al., 2021). Similarly, subjective perception has been shown to outweigh objective quality in many cases, particularly in pedestrian and cycling environments, making it essential to consider both dimensions of the data when evaluating multimodal integration (Rodriguez-Valencia et al., 2020; Calvey et al., 2015; Bellizzi et al., 2021).

The implications are far-reaching. Public Administrations, transport companies and operators, and urban designers can use quality assessments not only to improve single-mode services and/or infrastructure, but also to align their efforts around user-centred and context-sensitive multimodal networks (Julio et al., 2022; Barabino, 2018). Moreover, differences in quality perception across demographic and socio-cultural groups emphasize the need for inclusive planning, that considers the diverse needs of users (Podgorniak-Krzykacz et al., 2022; Rosa & Da Cruz Lopes, 2019; Ujjwall & Bandyopadhyaya, 2021).

Looking ahead, future research should aim to develop integrated and scalable multimodal quality frameworks that bridge the gap between system-level operational metrics and individual user experiences. There is a growing need for cross-platform tools that combine real-time data (e.g., GPS tracks, mobile app usage, service availability) with subjective feedback collected through surveys or digital interactions (Shin, 2020; Arellana et al., 2019). Moreover, cross-contextual studies are required to validate quality parameters across cities and populations of different contexts, ensuring generalizability and responsiveness to changing mobility habits (Larranaga et al., 2018; Da Rocha et al., 2019). Lastly, the increasing role of technology, such as journey planning apps, demands that future quality models incorporate digital experience as an inherent part of the multimodal mobility system (Liang et al., 2021; Hsu et al., 2018). Based on this evidence, assessing the quality of multimodal transport systems is critical not only to improve individual components, but also to build coherent, inclusive, and resilient urban mobility networks that effectively respond to user needs and urban challenges.

In summary, the rationale for assessing transport quality varies across modes but converges around a common purpose: improving user experience, promoting sustainability, and supporting evidence-based mobility planning. Despite the progress achieved, the literature still reveals significant fragmentation, with most studies focusing on single-mode evaluations and limited attention given to the interactions among modes. The growing importance of multimodal mobility highlights the need for comprehensive regulatory frameworks that integrate objective quality parameters and subjective user perceptions across all systems. This recognition motivates the next stage of the analysis, which examines the methodological approaches and techniques employed in transport quality assessment, setting the foundation for the selection of KQPs developed in the following chapters.

2.2.4 Methods and techniques for measuring quality

Building upon the rationale for assessing transport quality discussed in the previous section, this part examines the methods and tools most frequently employed in the scientific literature. It aims to define how the quality of sustainable and multimodal transport systems is measured and to identify emerging methodological trends.

The review of 49 selected papers reveals a wide diversity of methodological approaches used to assess the quality of sustainable transport systems (*Table 4*).

Regarding data collection, most studies rely on primary data obtained through surveys, questionnaires, and interviews. Satisfaction surveys are particularly widespread, serving as the backbone for user-centric evaluations (e.g., Xue et al., 2022; Julio et al., 2022). Other studies integrate observational techniques (e.g., Vujičić & Prester, 2019; Ismail et al., 2020), and novel sources like social media reviews (Shin, 2020). A growing interest is also noted in using GIS-based data (e.g., D'Orso & Migliore, 2019) and street-level imagery (Du & Huang, 2022) for spatial analysis.

The most commonly used software tools include SPSS, MATLAB, R, and GIS platforms. These tools support both statistical and spatial analysis. In studies with more advanced modelling needs, tools like AMOS, LINGO, DepthmapX, and machine learning frameworks, such as the Random Forest used by Du & Huang (2022) are adopted to process complex or large-scale data.

In terms of analytical techniques, a significant portion of the literature relies on multivariate statistical methods, including Principal Component Analysis (PCA), Exploratory Factor Analysis (EFA), and Confirmatory Factor Analysis (CFA). Structural Equation Modelling (SEM) is frequently used to explore relationships between latent variables (e.g., Chauhan et al., 2021; Vallejo-Borda et al., 2020). For decision-making under uncertainty, several studies utilize Multi-Criteria Decision Analysis (MCDA) methods such as DEMATEL, ANP, VIKOR, and AHP (e.g., Ma et al., 2019; Hsu et al., 2018). These are often coupled with fuzzy logic to handle data inputs.

Focusing on the method types used in this field of research, the analysis reveals that 35% are based on descriptive analyses, 16% on inferential analyses, and the remaining 49% adopt a mixed-methods framework that combines both descriptive and inferential techniques.

This distribution suggests a clear preference for integrating different levels of analysis, with nearly half of the papers aiming to both characterize patterns and statistically test relationships or hypotheses. While descriptive methods remain foundational for capturing user perceptions and contextual variables, the widespread adoption of combined approaches reflects an increasing recognition of the need for methodological robustness. This trend underscores the field's increasing commitment to generating evidence-based insights that are both empirically grounded and statistically validated, which is essential for evaluating and enhancing the quality of sustainable transport systems.

Table 4. Data collection, tools, method, and analysis used to assess the quality of urban and sustainable transport systems. The papers are listed in chronological order, from the earliest to the most recent.

Main sources	Mode[1]	Data collection[2]	Tools[3]	Method[4]	Analysis type[5]
Sarker, 2003	PM	MQMs	GIS	Qualitative exploration, drawing on existing literature along with examples of comfortable pedestrian space	D
Tan et al., 2007	PM	S-BMs	SPSS	Surveys on 12 road segments (pedestrian LOS + traffic volume), SPSS analysis, Pearson correlation, Stepwise regression modelling	D/I
Hsu & Lin, 2011	CM	S-BMs	GIS	Multicriteria assessment, Fuzzy AHP for weighting, Floyd-Warshall shortest path, Five-step bicycle suitability model	D/I
Amoroso et al., 2012	PM	SDS (literature review)	ns*	Pedestrian mobility indicators, Walking audits (PERS, PEQI), Index validation, Accuracy & adaptability assessment	D
Ely et al., 2012	LPT	O-BMs & S-BMs	SketchUp©, Excel©	3D modelling, Attributes Grid (MEGA)	D
Argin & Ozbil, 2015	PM	S-BMs & I-BMs	GIS + DedepthmapX	Angular segment & connectivity analysis, Linear & regression models, Street environment audits	D/I
Calvey et al., 2015	CM	S-BMs & MQMs	ns*	EFA, Field testing with instrumented bicycle (vibration data), Analysis of vegetation, surface quality/issues, night lighting	D/I
Douglas, 2015	MMT	S-BMs	ns*	Combined Stated Preference + Rating Approach	I
Talavera-Garcia & Soria-Lara, 2015	MMT	S-BMs	GIS	Indicator selection & threshold setting, Web survey for weighting factors, Q-PLOS formula application, Comparative analysis in two mobility contexts	D/I
Blečić et al., 2016	PM	MQMs (CFS)	ns*	Regression analysis, Data aggregation by street segment	D/I
Diana et al., 2016	MMT	S-BMs	ns*	PCA	D/I
Beura et al., 2017	PM	S-BMs	HD video cameras	Stepwise multivariable regression; HD video traffic flow recording; Fuzzy C-Means clustering; Pearson correlation; t-test significance	D/I
Diana et al., 2017	MMT	S-BMs	ns*	Exploratory & data-driven analysis, Satisfaction classification, Cronbach's alpha, PCA with varimax rotation	D/I
Podciborski, 2017	LPT	S-BMs	ns*	Scaling technique that relies on a point-based system	D
Soltani et al., 2017	PM	S-BMs	ns*	Factor analysis, non-parametric Friedman ranking test	D/I
Barabino, 2018	LPT	OPMs (Mystery Client, AVL)	Excel©, Spreadsheets	HAN+OUT algorithm	D
Groenendijk et al., 2018	MMT	S-BMs	ns*	BWM	D
Hsu et al., 2018	BSS	S-BMs	ns*	DEMATEL for causal modelling, ANP (DANP) for weighting, VIKOR for decision-making under uncertainty, Grey Theory for handling imprecise data	I
Larranaga et al., 2018	PM	S-BMs & I-BMs	GIS, NGene software, Google Street View	Profile case analysis, D-efficient experimental design, Random utility model, Walkability scores from discrete choice models	D/I
Arellana et al., 2019	PM	S-BMs	GIS, Google Street View	Factor selection, Weight estimation via rank perception survey & discrete choice model, Data collection & walkability index, MNL model	D/I
Bellizzi et al., 2019	PM	S-BMs	ns*	Experimental data collection & statistical analysis	D
Da Rocha et al., 2019	PM	S-BMs	ns*	SQI calculation combining technical and perceptual data	D
D'Orso & Migliore, 2019	MMT	SDS	GIS + Google Street View	Indicator scoring and weighting, quality index calculation, priority index formula, thematic mapping	D
Ma et al., 2019	BSS	S-BMs	ns*	DEMATEL for criteria weighting and interdependencies, VIKOR for stakeholder-based ranking, questionnaire based on revised SERVPERF model	I

Main sources	Mode[1]	Data collection[2]	Tools[3]	Method[4]	Analysis type[5]
Rosa & Da Cruz Lopes, 2019	LPT	S-BMs	SPSS	Data analysis	D
Rowangould & Corning-Padilla, 2019	PM	I-BMs	GIS	Focus groups, Linear regression analysis	D/I
Sousa et al., 2019	PM	CMs (theoretical exposition + pilot survey)	GIS	Survey design & team training, Data collection, PEQI calculation, GIS-based visualization & mapping	D
Vujičić & Prester, 2019	LPT	O-BMs	Check-list	Modified P-TRANSQUAL model, control chart, regression analysis	D/I
Fistola et al., 2020	CM	SDS (literature review)	Mendeley, EndNote, Textstat, GIS, Google Earth	Equivalent length metric using mathematical equation	D
Ismail et al., 2020	PM	S-BMs & O-BMs	SPSS	Mixed methods (observation, photos, survey with 100 respondents), SPSS analysis, Factor analysis (3 dimensions), Validation with KMO & Bartlett's tests	D/I
Jahan et al., 2020	PM	S-BMs	ns*	PCA, factor analysis, SEM	D/I
Raswol, 2020	PM	S-BM & O-BMs	Checklist	Qualitative assessment based on onsite observation, survey, student questionnaire, walkability checklist	D
Rodriguez-Valencia et al., 2020	PM	S-BMs & O-BMs	ns*	Random intercept surveys (perception + objective data), OLS regression with geometric, operational & perceptual variables, Model fit via R^2 and F-tests	D/I
Rossetti, 2020	LPT	O-BMs	GIS	Indexes calculation, GIS-based mapping and layer analysis	D
Shin, 2020	BSS	SDS (social media)	ns*	Descriptive analysis of numeric ratings, Content analysis and manual coding of narrative reviews	D
Vallejo-Borda et al., 2020	PM	S-BMs	ns*	EFA, SEM (max likelihood), scale-based latent variable modelling	I
Vallejo-Borda et al., 2020	PM	S-BMs	ns*	Sidewalk measurements, Ordered Probit MIMIC model, Model validation (match score, error distribution, chi-squared), Split-sample validation	D/I
Bellizzi et al., 2021	PM	S-BMs	ns*	Descriptive statistics, MNL model, discrete choice modelling	D/I
Burlando et al., 2021	MMT	S-BMs	CATI	Quantitative non-aggregative method, POSet	D
Chauhan et al., 2021	MMT	S-BMs	ns*	EFA, CFA, SEM, three-stage factor selection, model validation with split samples	I
Liang et al., 2021	BSS	I-BMs & S-BM	LINGO, MATLAB	Fuzzy BWM for criteria weighting, VIKOR for alternative ranking, fuzzy logic with linguistic variables to manage uncertainty	I
Ujjwal & Bandyopadhyaya, 2021	PM	S-BMs	ns*	PCA, D-Optimal design, OPM, Model validation via actual vs. predicted ratings	D/I
Ujjwal & Bandyopadhyaya, 2021	PM	S-BMs	ns*	EFA for parameter clustering, Constant Sum Paired Comparison, PLOS model with perception ratings & weightages	D/I
de Aquino Traldi et al., 2022	CM	I-BMs	Microsoft Excel software spreadsheets, Google Earth Pro, GIS	AHP for cycling infrastructure evaluation, Geostatistical modelling (variogram, ordinary kriging), Cross-validation, GIS mapping	D/I
Du & Huang, 2022	PM	S-BM & I-BMs	Streetscape images	PSPNet for semantic segmentation, Regression analysis, Random Forest for large-scale prediction, Pre-processing of streetscape features	D/I
Hamerska et al., 2022	E-MM	S-BMs (CAWI)	R programme	MMQUAL model, CFA	I
Julio et al., 2022	BSS	S-BMs	STATA 15, MATLAB	Combined IPA-3FT model	D
Podgorniak-Krzykacz et al., 2022	BSS	I-BMs (CAPI)	IBM SPSS Statistics 25.0	EFA (PCA, Oblimin rotation), KMO and Bartlett's test for validation	I
Xue et al., 2022	BSS	S-BMs	SPSS, AMOS	ANOVA, PCA, Reliability tests (Cronbach's alpha, Composite Reliability), SEM, AVE	D/I

Main sources	Mode[1]	Data collection[2]	Tools[3]	Method[4]	Analysis type[5]

[1] Transport Modes:

BSS: Bike-Sharing System
CM: Cycling Mobility
CMs: Combined methods
I-BMs: Interview-Based Methods
LPT: Local Public Transport
MMT: Multimodal Transport
MQMs: Mixed and Qualitative Methods
O-BMs: Observation-Based Methods
OPMs: Objective Performance Measures

PM: Pedestrian Mobility
S-BMs: Survey-Based Methods
SDS: Secondary data sources

[2] Data collection:

AVL: Automatic Vehicle Location
CAPI: Computer-Assisted Personal Interviewing
CAPI: Computer-Assisted Personal Interviewing
CATI: Computer-assisted telephone interviews
CAWI: Computer Assisted Web Interviewing
CFS: Contingent Field Survey

[3] Tools:

AMOS: Analysis of Moment Structures
GIS: Geographic Information System
LINGO: LINear Generalized Optimization
ns*: none specified
MATLAB: MATrix LABoratory
SPSS: Statistical Package for Social Science or Statistical
STATA: Statistics and Data

[4] Method:

AHP: Analytic Hierarchy Progress
ANOVA: Analysis of Variance
ANP: Analytic Network Process
AVE: Average Variance Extracted
BWM: Best-Worst Method
CFA: Confirmatory Factor Analysis
CFS: Contingent Field Survey
DANP: DEMATEL-based Analytic Network Process
DEMATEL: DEcision MAking Trial and Evaluation Laboratory
EFA: Exploratory factor analysis
HAN: Handling algorithm
IPA: Importance-Performance Analysis
KMO: Kaiser-Meyer-Olkin
MEGA: Multi-faceted Evaluation of Grid Attributes
MIMIC: Multiple Indicators Multiple Causes
MMQUAL: MicroMobility QUALity
MNL: Multinomial Logit
OLS: Ordinary Least Squares
OPM: Ordered Probit Models
OUT: Output algorithm
PCA: Principal Component Analysis
PEQI: Pedestrian Environmental Quality Index
PERS: Pedestrian Environment Review System
PLOS: Pedestrian Level of Service
POSet: Partially Ordered Set
PSPNet: Pyramid Scene Parsing Network
PSS: Product and Service Solutions
P-TRANSQUAL: Public Transport Service Quality

Q-PLOS: Quality of Pedestrian Level of Service
SEM: Structural Equation Modelling
SERVPERF: Service Performance
SQI: Service Quality Index
VIKOR: VIseKriterijumska Optimizacija I Kompromisno Resenje [Multi-criteria Optimization and Compromise Solution]
3FT: Three-Factor Theory

[5] Analysis type:
D = Descriptive
I = Inferential

Overall, the methodological landscape shows a growing commitment to the integration of quantitative rigor and user-centred perspectives. This evolution reflects a shift from isolated measurements toward comprehensive evaluation frameworks capable of linking analytical techniques with meaningful quality parameters. This is a connection explored in the next section.

2.2.5 Quality criteria and parameters in the literature

This section concerns the main quality criteria and parameters used by experts and seeks to define what is measured when assessing the quality of sustainable transport systems. This literature review identified the main quality criteria and parameters used to evaluate pedestrian, bicycle, LPT, and electric micromobility systems.

The review yielded a wide range of parameters with varying levels of specificity. To organize them systematically, a two-level hierarchical classification was adopted: broad thematic categories (first-level parameters) and their more detailed subcomponents (second-level parameters). Terminological inconsistencies across sources were resolved through a process of normalization, ensuring clarity and consistency in parameter naming. This process led to the identification of 23 first-level and 65 second-level parameters.

A higher-level grouping was then introduced, based on the quality criteria defined by European standards for public transport service quality - EN 13816:2002 and EN 15140:2006. These standards establish eight core quality criteria: accessibility, availability, comfort, customer care, environmental impact, information, safety and security, and time. These criteria reflect both regulatory expectations and user-centered objectives commonly adopted in transport planning and service assessment.

Each parameter was associated with one of these overarching quality dimensions, based on its thematic relevance and the definitions provided by the standards, thus providing a coherent framework for analysis and comparison. *Table 5* presents the resulting hierarchy of quality criteria and parameters.

Accessibility includes parameters related to the ease of reaching and using a transport system. This criterion includes physical access to infrastructure, such as docking stations in bike share systems (Shin, 2020), as well as solutions for specific user groups, such as ramps for passengers with strollers or wheelchairs (Podciborski, 2017). It also refers to the system's ability to connect with other modes of transport (Rossetti, 2020; Podgorniak-Krzykacz et al., 2022), highlighting its inherently multimodal character. Availability refers to the extent and continuity of service, including spatial coverage (e.g., areas served by bike share systems) (Julio et al., 2022), opening hours (Liang et al., 2021), and frequency of service (Arellana et al., 2019). The parameters that fall under this quality criterion apply broadly to all transport systems. Comfort refers to the factors that influence the pleasantness of travel. These include on-board and in-station services (Rosa & Da Cruz Lopes, 2019; Vallejo-Borda et al., 2020), infrastructure design (Sarker, 2003; Argin and

Ozbil, 2015; de Aquino Traldi et al., 2022), functionality within urban space (Amoroso et al., 2012), and waiting conditions (Ely et al., 2012; Vujičić & Prester, 2019). Environmental aspects such as cleanliness (Jahan et al., 2020; Ismail et al., 2020; Ujjwal and Bandyopadhyaya, 2021), maintenance (Blečić et al., 2016; Raswol, 2020) and the presence of amenities, landmarks or natural scenery (Da Rocha et al., 2019) also contribute. Given its breadth, comfort is the most frequently cited criterion and is adaptable to all transport modes, although not uniformly across all settings. Customer service includes parameters related to user support, staff interaction, ticketing, and usability of user interfaces (Hsu et al., 2018; Ma et al., 2019; Xue et al., 2022). This dimension is particularly relevant to services of LPT, bike sharing, and micromobility but less applicable to pedestrian mobility. Environmental impact focuses on the ecological consequences of transportation, including air pollution (Calvey et al., 2015; Rodriguez-Valencia et al., 2020; Podgorniak-Krzykacz et al., 2022), resource consumption, and land use (Sousa et al., 2019). Information refers to the availability and quality of information that supports user understanding of other quality criteria, such as security (Barabino, 2018), accessibility (de Aquino Traldi et al., 2022), and others. Safety and security include measures to protect users from accidents (safety) (Tan et al., 2007; Beura et al., 2017; Larranaga, 2018; Bellizzi et al., 2021; Du and Huang, 2022) and from crime or harmful behaviour (security) (Soltani et al., 2017; Rowangould and Corning-Padilla, 2019; Bellizzi et al., 2019; Fistola, 2020). Finally, time is concerned with aspects of travel planning and execution, such as travel time (e.g., time spent walking to a destination) (Ismail et al., 2020), as well as adherence to schedules, reliability, and punctuality (Diana et al., 2016, 2017).

As this review demonstrates, the assessment of transport quality is a complex and multidimensional topic, involving a wide range of parameters that reflect different user needs and the characteristics of multiple transport systems. Quality criteria and related parameters may relate to infrastructure or service aspects, be more or less relevant to different transport systems.

Table 5. Criteria and parameters to assess the quality of urban and sustainable transport systems.

Criteria	**Parameters – level I**	**Parameters – level II**
Time	Adherence to schedule / Reliability	Efficiency
		Punctuality
		Regularity
	Length of trip time	At Origin/boarding & alighting/Destination points
		Time in-vehicle
		Time for trip planning
Safety & Security	Emergency management	Facilities and plans
	Freedom from accidents	Crashes
		Avoidance/visibility of hazards
		Travel behaviour
		Presence/visibility of supports, e.g. handrails
		Preventive design
		Speed
		Traffic flow
	Freedom from crime	Lighting
		People (mis)behaviour
		Public safety monitoring

Information	General information	Info. about accessibility
		Info. about availability
		Info. about security
		Info. about t source of information
		Info. about time
		Info. about comfort
		Info. about environmental impact
		Info. about customer care
Environmental impact	Pollution and health	Exhaust
		Natural resources
Customer care	Assistance	Customer orientation
	Commitment	Innovation and initiative
	Customer interface	Complaints
		Tools and related qualities
	Staff	Staff appearance
		Staff availability
		Staff skills (behaviour)
	Ticketing options	Fare structure and policy
		Consistent price calculations
Comfort	Ambient conditions	Atmosphere
		Cleanliness
		Disruption
		Density
		Visual pollution
		Odour
		Noise
		Weather protection
	Complementary facilities	Commercial activities and services
		Environment facilities
		Entertainment
		Amenities
	Ergonomics	Ease of movement
		Furniture design
	Ride comfort	Driving
	Path comfort	Walking and cycling
	Seating and personal space for public transport vehicles	In-vehicle
		At b/a points
Availability	Network	Modes
		Area covered
		Distance to O/D points
	Operation	Frequency
		Operating hours
		Load factor
Accessibility	External interface	Interchange/multimodality
	Internal interface	Entrance/exits
		Internal movement
	Ticketing availability	Ticketing acquisition
		Ticketing validation

Although no universally accepted taxonomy was identified in the literature, several recurring thematic clusters emerged. These were consolidated into six overarching clusters: environmental and aesthetic (Calvey et al., 2015), functional design and technological (Amoroso et al., 2012), geometric and physical (Soltani et al., 2017), maintenance and conservation (de Aquino Traldi et al., 2022), operational (Rodriguez-Valencia et al., 2020), and social and psychological (Ely et al., 2012; Soltani et al., 2017). These clusters help contextualize each parameter according to the thematic field that influence system quality.

As with the criteria, the association between individual parameters and the identified clusters was carried out manually through a reasoned interpretive process. Where possible, previously established associations found in the literature were used as guiding

references. For example, if a study classified ‘sidewalk width’ under a “geometric” cluster, similar parameters such as ‘size’ or ‘spatial dimensions’ found in other studies were also grouped under the geometric/physical cluster. A similar approach was followed when attributing parameters to the EN-based quality criteria. This dual classification method enables a comprehensive, multidimensional understanding of transport system quality, aligning regulatory standards with practical and contextual insights from the literature.

Each quality criterion may encompass a combination of parameters related to several of these thematic clusters, reflecting the multifaceted nature of quality in transport systems.

Figure 16 presents a relational matrix between criteria and clusters, offering insights into how quality is conceptualized in the existing literature. The circles indicate the percentage of parameters associated with a given quality criterion that corresponds to a specific cluster. The total of all thematic clusters for each criterion equals 100%. For instance, the first row of the matrix shows that the “accessibility” criterion is predominantly composed of geometric and physical (62%) and operational parameters (35%), with smaller contributions from functional design and technological, and social and psychological dimensions (2%). Some criteria exhibit strong one-to-one associations, suggesting narrowly defined understandings of these relationships. For example, environmental impact (100%) is exclusively linked to environmental and aesthetic clusters. This indicates that environmental impact is predominantly framed in terms of sustainability and ecological or visual footprint, with minimal consideration of operational or technological aspects. Similarly, time (100%) is entirely associated with operational clusters, reflecting an understanding of time as a function of system characteristics such as travel time, delays, and service frequency. Information (98%) is almost entirely attributed to functional and technological design, underscoring the importance of real-time data, digital interfaces, and signage - consistent with the growing role of digital technologies in contemporary transport systems. Other criteria are more complex. Customer care, for instance, is primarily linked to social and psychological clusters (74%), with additional contributions from functional design (15%) and operational aspects (12%). This highlights the importance of user experience, interaction, and perception in defining service quality. Comfort emerges as the most multidimensional criterion, incorporating contributions from geometric and physical (38%), environmental and aesthetic (19%), social and psychological (18%), functional (11%), and maintenance (14%) clusters. This diversity reflects the wide range of elements influencing perceived comfort, from spatial characteristics and cleanliness to ambient conditions and design quality. Safety and security also span several cluster types, with substantial contributions from operational (39%), geometric and physical (27%), and social and psychological (31%) clusters. This pattern underscores the interplay between physical infrastructure, system operations, and user perception in shaping both objective and subjective safety. Finally, accessibility and availability show a strong reliance on geometric and physical (62% and 49%) and operational (35% and 51%) clusters, respectively. This reflects conventional approaches in transport planning that prioritise spatial design and service provision. The minimal involvement of

social/psychological and maintenance clusters suggests a potential gap in current frameworks, where aspects such as user inclusivity, adaptability, and infrastructure upkeep may be underexplored.

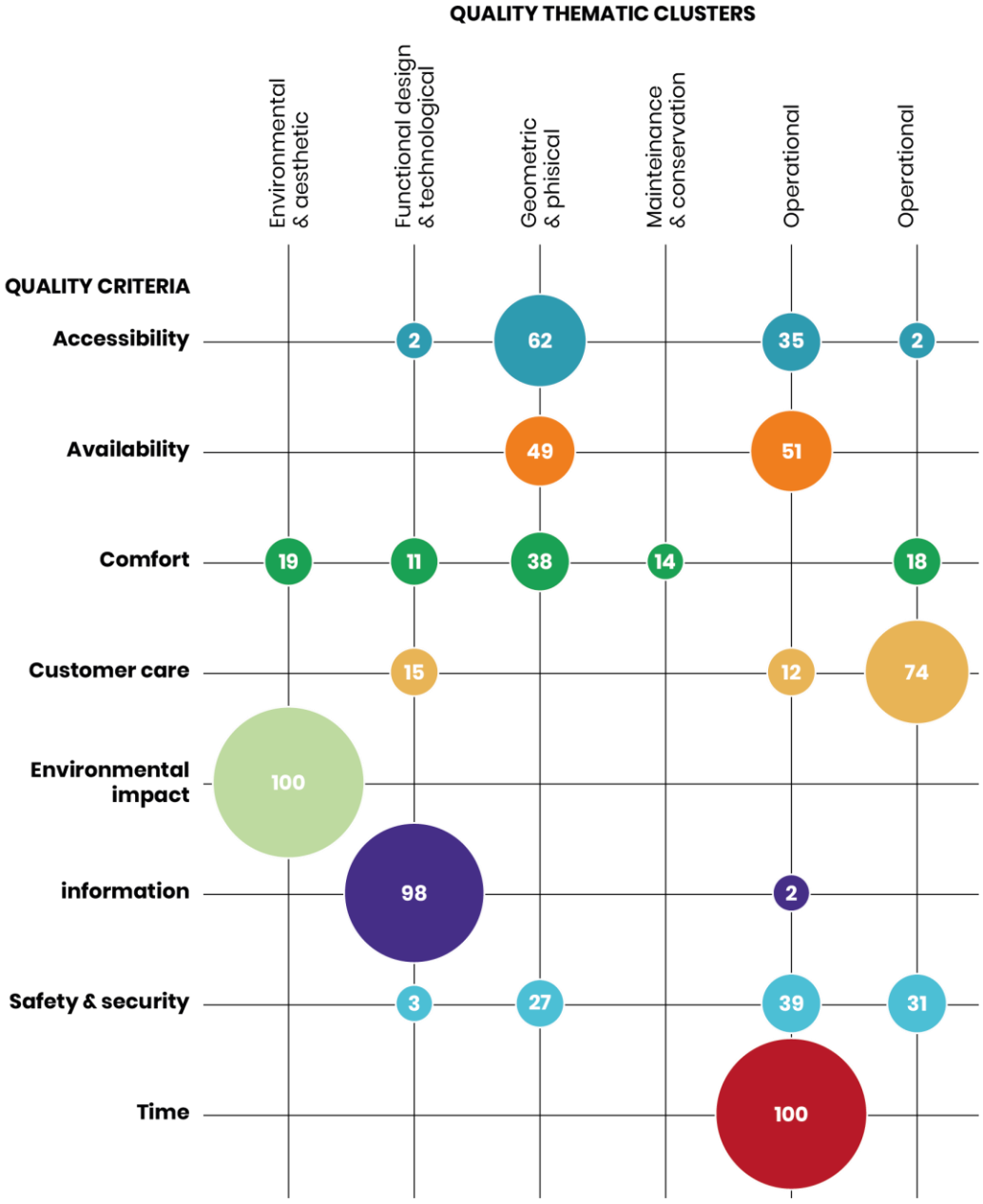


Figure 16. Summary matrix on the percentage of thematic clusters characterizing the eight quality criteria.

2.3 Discussion and identification of research gaps

The systematic analysis of transport quality presented in this chapter confirms that, although the concept of transport quality has evolved significantly, the research landscape remains highly fragmented. Most studies continue to analyse transport modes separately, with limited integration across systems and analytical dimensions. Even

frameworks such as LOS, QoS, SERVQUAL, PLOS/QLPOS, MMLOS, and emerging multimodality-oriented tools illustrate a clear progression, yet remain sectoral and insufficiently integrated. As multimodality becomes a priority in urban policy, this lack of integration limits the ability to evaluate how different transport modes interact and contribute to the overall travel experience. This fragmentation stems from the lack of a shared conceptual basis for understanding multimodal quality.

Indeed, a first major gap emerging from the literature concerns two crucial aspects. First, the absence of a universally accepted definition of multimodal quality that reflects current sustainability goals, integrating the perspectives of all actors involved in the provision, management, and use of transport systems, and accounting for the full range of factors that shape user perceptions. Such a definition must extend beyond operational or infrastructural aspects to include social, environmental, technological, and experiential spheres, and be representative of the different dimensions of quality (designed, delivered, expected, and perceived). Secondly, the absence of a structured and shared evaluation framework for assessing the quality of multimodal transport systems that meet the prerequisites just expressed.

Methodologically, the field has shifted from single-mode, technically oriented evaluations to more comprehensive, hybrid approaches. Multi-criteria decision analysis (MCDA), structural modelling, GIS-based analysis, and increasingly data-driven techniques (e.g., machine learning) are now frequently used to address the complexity of urban mobility and its quality. This complexity also emerges from the large number of quality parameters and indicators found in the literature.

The proliferation of criteria, parameters, and indicators across studies highlights the absence of a standardised and transferable framework for prioritising the most relevant quality dimensions, particularly in multimodal and sustainability-oriented contexts. This represents another significant gap. The literature shows an extremely heterogeneous landscape, in which a lot of quality parameters are proposed, often differing across modes, geographical contexts, and methodological traditions. While this richness reflects the multidimensional nature of transport quality, it also generates ambiguity and inconsistency, making it difficult to determine which parameters are essential, comparable, or sufficiently representative of system performance.

Moreover, no widely accepted method exists for selecting or filtering these parameters, nor for assessing their relative importance in multimodal settings. As a result, quality assessments often rely on excessively long lists of indicators, increasing data requirements and operational complexity, or alternatively use overly narrow sets of parameters that fail to capture the full travel experience. A structured, evidence-based approach to parameter selection is therefore crucial to support robust, scalable, and policy-relevant multimodal quality assessment.

Moreover, rarely percentage-based metrics to assess multimodal quality from the user's perspective are included in the evaluation methodology. Specifically, it refers to the

percentage of users who expect a predefined level of service; for whom the service is designed; who receive the planned service; and who perceive the service as actually delivered. Overall, these metrics can be used to manage quality, identify gaps, and suggest corrective measures, emphasising a user-centric approach to multimodal quality assessment.

At the same time, the literature reveals important signs of convergence. Transport quality is consistently described as a multidimensional construct, encompassing both tangible (infrastructure, operational performance) and intangible components (user perception, social factors, psychological comfort). Across studies, a coherent set of recurrent criteria emerges - accessibility, availability, comfort, customer care, environmental impact, information, safety & security, and time - largely derived from European standards and widely applied across transport sectors. These criteria represent important cornerstones to be included in the structure of a shared and robust quality assessment framework.

3. THE NEW CONCEPT OF MULTIMODAL TRANSPORT QUALITY

Based on the previous literature review and in contrast to the traditional concept of QoS, this chapter proposes an approach to multimodal quality assessment, which introduces a new definition of multimodal transport quality applicable to an integrated and sustainable urban mobility scenario *Figure 17*.

In a multimodal transport system, QoS must be extended to multiple systems in addition to LPT and the entire travel experience, where different modes of transport are included within the same travel chain. For instance, in a simple case, an overall QoS assessment should consider the quality of the pedestrian path a user covers to reach a bus stop, the quality of the bus stop area, the quality of the ride on board, and the quality of the pedestrian path to reach the final destination. In a trickier case, the (multimodal) journey chain may also cover stretches of travel made by bike-sharing services, private bicycles, or electric scooters. It is evident that by increasing the modes of transport, the stakeholders involved and the viewpoints to be considered in the assessment also increase. Considering the interaction between these factors, the proposed approach looks to the full range of characteristics necessary for a sustainable urban transport system, capable of improving quality and addressing the critical challenges faced by contemporary cities, including ecological transition and urban regeneration.

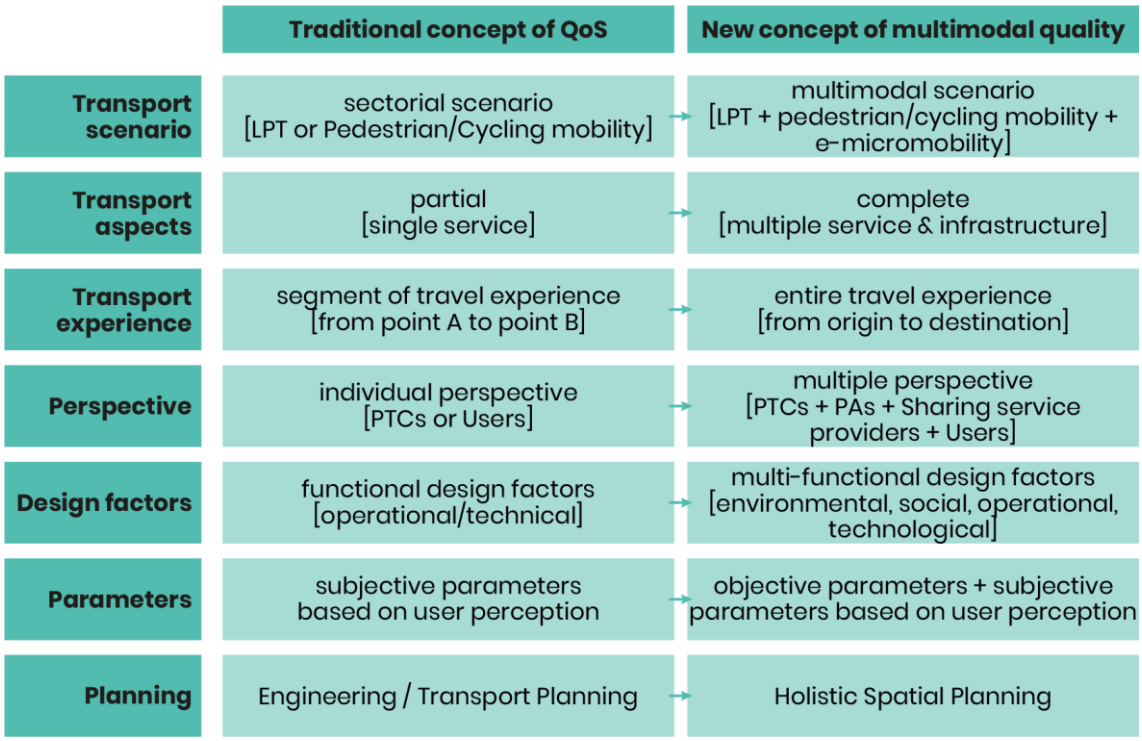


Figure 117. A summary of the change in perspective from the traditional concept of Quality of Service to the new concept of Quality of a Multimodal and Sustainable Urban Transport System.

Multimodal transport quality can be defined as the capacity of an integrated, multi-modal transport system - encompassing various sustainable travel modes - that fully aligns with both current and potential users' needs throughout the entire travel experience. This system delivers high-performing infrastructure and services that collectively support a seamless, efficient, and inclusive travel experience across all journey stages.

This concept extends beyond operational performance to encompass multifunctional design dimensions, including technical, social, environmental, and technological factors. It adopts a holistic perspective that integrates the objectives and responsibilities of transport providers, the regulatory and planning roles of public authorities, and the diverse needs, expectations, and behavioural patterns of users. Particularly, users' perspectives are fundamental in assessing the quality of a transport system, as they serve as direct evaluators of its quality. However, in the context of multimodal systems, the assessment process is complicated by the multiplicity of aspects arising from the presence of several systems and users' perceptions may be subject to greater variability. As the different system attributes, attitudes toward transport modes, and individual preferences can influence the assessment, subjective evaluations alone may not provide a comprehensive or accurate measure of service quality. To obtain a more robust and reliable assessment, it is therefore essential to complement subjective parameters based on users' perceptions with objective measurable on-site parameters.

3.1 Dimensions and framework for assessing transport quality

Building upon the proposed conceptualization of multimodal quality, this study adopts the Logical Framework Approach (LFA) to structure the quality conceptual framework from a result-oriented perspective (USAID, 1969). Originating in project and program management, the LFA has recently been adapted to transportation planning as a mean to align inputs (tools and resources), outputs (direct results), outcomes (expected results at the system level), and the sustainability-oriented goals as long-term objectives that the framework aims to support (Russo & Rindone, 2023). In this context, it enables the structured identification of problems and opportunities, which can be framed as quality gaps in the transport system, and supports the formulation of targeted, context-specific interventions consistent with clearly defined objectives and expected results. The use of Objectively Verifiable Indicators (OVIs) and clearly defined means of verification enhances both methodological transparency and replicability across different urban contexts. This study extends the LFA's application from the planning of sector-specific transport services to the broader and more complex domain of urban multimodal mobility. Such an extension emphasizes the need for structured approaches capable of capturing the interdependencies among diverse transport modes, actors, and quality dimensions. The structured matrix developed within this framework provides a standardized and transferable tool for assessing multimodal transport quality.

To operate this conceptualization, the proposed conceptual framework consists of three interrelated phases, each corresponding to a specific step in the assessment and management of quality within sustainable and multimodal transport systems. The phases are aimed at supporting a systematic, actionable process for enhancing urban mobility quality across all modes, including pedestrian, bicycle, micro-electric mobility, and public transport.

The first phase concerns the characterization of the system under investigation. Given the complexity and multidimensional nature of transport system quality, this phase focuses on identifying and selecting the key quality characteristics that define a sustainable and multimodal urban transport system. Each aspect of a transport system constitutes a broad analytical domain encompassing a set of criteria, parameters, and indicators. These components form the essential references for the subsequent assessment processes, ensuring that both service and infrastructure aspects are adequately considered. Defining these elements at the outset is crucial to allow for consistent monitoring practices and meaningful quality assessments, ensuring that the assessment process yields meaningful insights (Barabino, 2018). This phase consists of a comprehensive literature review and a multicriteria analysis aimed at selecting the most relevant quality parameters, particularly in cases where the literature yields an excessively large number of indicators to be effectively managed within the proposed evaluation framework.

The second phase focuses on the measurement of system quality. It integrates both objective and subjective methods to provide a balanced and comprehensive assessment. Objective data are collected through field observations, and quantitative indicators. Differently, subjective data are gathered through user surveys, which capture perceptions, expectations, and experiences across the different mobility modes considered. The combination of these data sources allows for a more nuanced understanding of system performance, highlighting areas where technical deficiencies align (or fail to align) with user satisfaction. The dual approach strengthens the reliability of the assessment by bridging measurable performance with perceived quality.

The third and final phase concerns the management of the transport system. It focuses on identifying quality gaps and formulating targeted improvement strategies. GIS-based spatial analyses play a relevant role in mapping weaknesses and contextualizing them spatially. The aim is to identify or design context-specific actions that address the shortcomings and contribute to enhancing the overall quality of urban transport services. These actions are intended not merely as operational adjustments but as part of a strategic framework that supports public policy objectives, promotes targeted interventions, and advances the transition towards more sustainable, equitable, and user-oriented mobility systems.

This structured, multi-phase approach is summarized in the logical framework below (*Figure 18*), which highlights the relationship between tools, results, and expected outcomes across each phase of the framework.

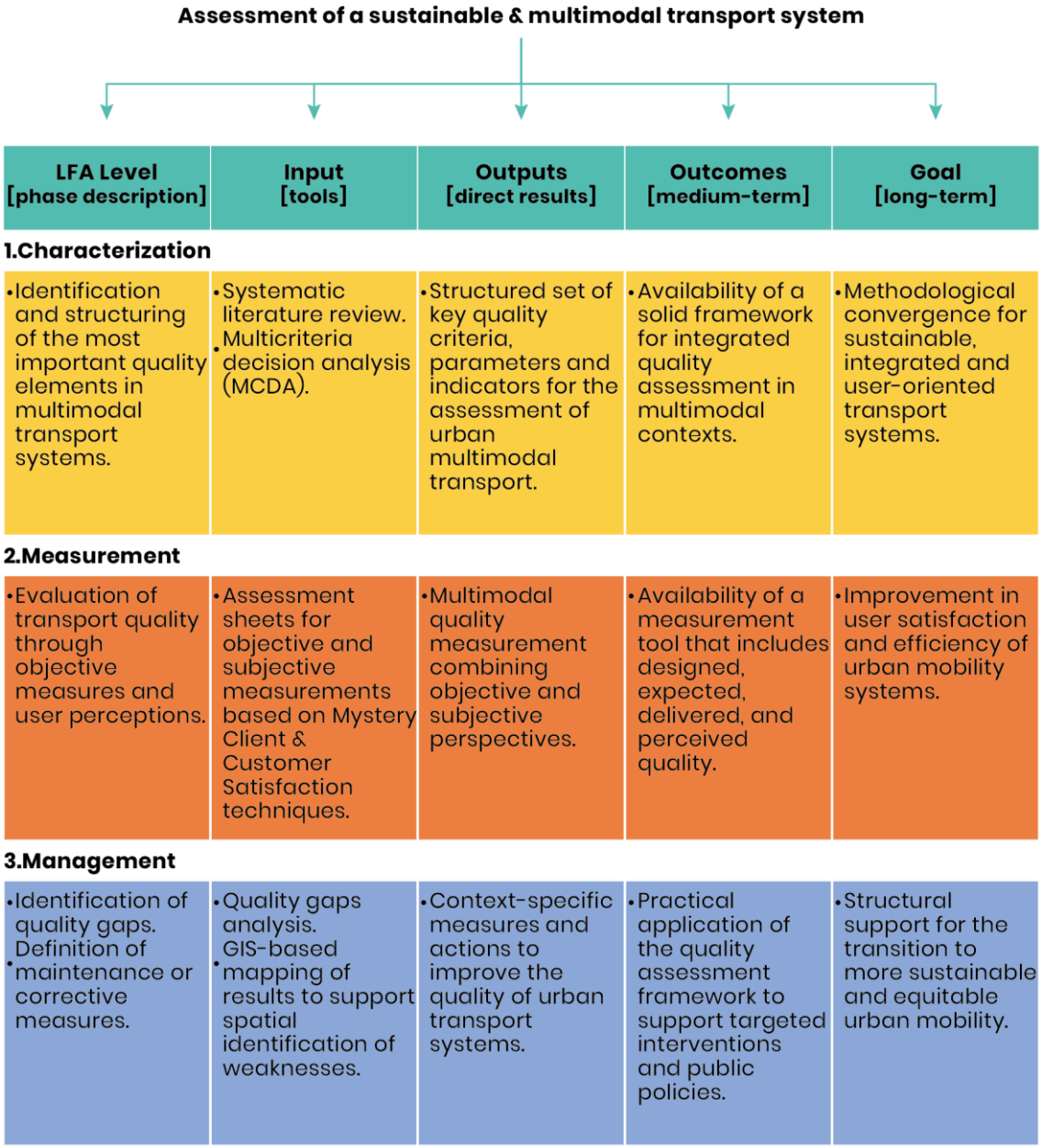


Figure 18. Relationship matrix among inputs, outputs, expected outcomes and goals across each phase of the conceptual framework proposed for the quality assessment of a multimodal and sustainable urban transport system, based on the Logical Framework Approach (LFA).

Among the methods emerging in literature, the MMLOS framework is among the most recognized methods for evaluating transport quality across multimodality. While MMLOS provides a valuable foundation, it is limited in scope as it excludes emerging micromobility modes and environmental quality aspects. Instead, recent frameworks like MMQUAL address the specificities of shared and e-micromobility systems. Additionally, other approaches, such as SERVQUAL, emphasize the subjective user experience and have been instrumental in shifting the focus toward passenger-centered quality assessments. Building on these existing methods and on their evolution over time, the present study adopts the LFA to structure a comprehensive, result-oriented framework that bridges

technical performance and user perceptions across all relevant transport modes. By incorporating both objective data and subjective user feedback, this framework extends prior work by explicitly integrating micromobility and sustainability dimensions. The LFA further supports systematic identification of quality gaps and targeted interventions, enhancing replicability across diverse urban contexts. In this way, this approach not only complements but also advances current quality assessment practices by providing a transferable conceptual framework aimed at supporting strategic planning and operational improvements in complex multimodal transport systems.

4. METHODOLOGY

This chapter presents the research methodological framework, structured into three main phases that reflect a progressive and interrelated process of quality characterization, measurement, and management within sustainable and multimodal urban transport systems (*Figure 19*). Each phase builds on the findings of the previous one, contributing to the development of a comprehensive and replicable approach to assessing and improve transport quality.

The first phase focuses on the characterization of transport quality and the selection of KQPs. Based on the results of the systematic literature review, this stage identifies the most relevant criteria and parameters of quality and applies multi-criteria decision analysis (MCDA) - specifically the Analytic Hierarchy Process (AHP) and Monte Carlo simulations - to prioritize and validate them.

The second phase concerns the measurement of transport system quality, operationalized through an integrated methodological framework that combines Mystery Client and Customer Satisfaction approaches. This phase enables a multi-perspective evaluation of the designed, expected, delivered, and perceived quality of service in a real urban context, linking objective observations with users' perceived experience.

The third phase addresses the management of transport quality, using the results from the measurement stage to identify quality gaps and define improvement strategies. It involves quality gaps analysis and the formulation of guidelines and recommendations to support decision-making and enhance the overall performance of urban mobility systems.

Overall, this chapter details the methodological approaches, tools, and data sources employed in each phase. The corresponding results are presented in Chapter 5, while their discussion and interpretation are developed in Chapter 6, ensuring a clear separation between methodological design and analytical outcomes.

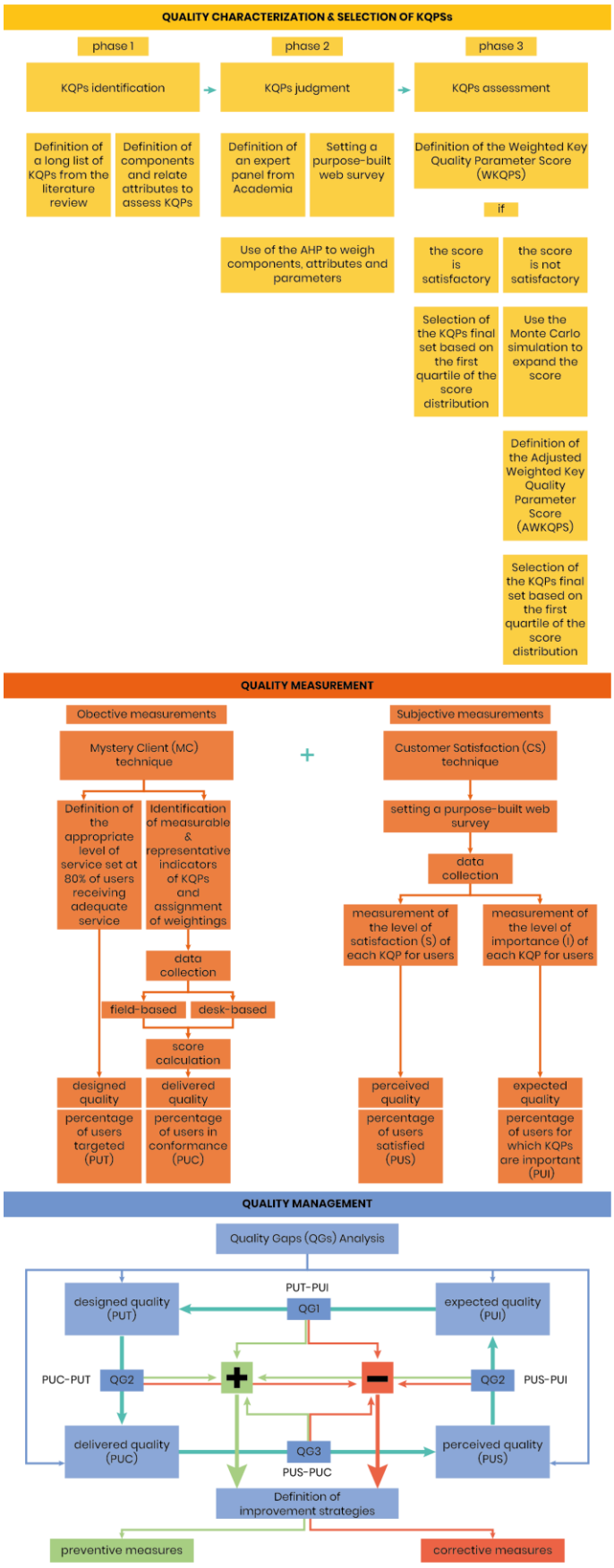


Figure 19. Summary of the three methodological phases of the research: characterisation and selection of KQPs, quality measurement and quality management.

4.1 Characterization of transport quality and Selection of Key Quality Parameters

The systematic literature review presented in the previous chapter highlighted the wide diversity of parameters used to define and assess transport quality. This multiplicity of parameters reflects the intrinsic complexity of the concept that characterizes the study of transport systems. However, it also poses a challenge for both research and practice: how to identify, among many possible parameters, those that are truly representative of the quality of sustainable and multimodal urban mobility systems.

The need to reduce the fragmentation of quality assessment frameworks and to promote a more integrated approach constitutes the rationale for this chapter. While early methodologies primarily focused on LPT and were designed to assist transport operators in monitoring service quality, the increasing emphasis on sustainability, user experience, and multimodal integration now requires broader frameworks capable of encompassing non-motorized mobility (NMM) and electric micromobility (e-MM) as integral components of urban transport systems.

The objective of this chapter is therefore to identify and prioritize a set of KQPs that can effectively represent the essential dimensions of quality in sustainable and multimodal urban transport. This process aims to move from the comprehensive but heterogeneous collection of parameters obtained through the literature review to a coherent and operational subset of variables suitable for comparative evaluation and policy application.

To achieve this, the methodology combines expert judgment, multicriteria decision analysis (Analytical Hierarchy Process - AHP), and Monte Carlo simulation. The integration of these approaches allows the inclusion of expert knowledge from different disciplines, the quantitative evaluation of the relative importance of each parameter, and the adjustment of results to account for uncertainty and bias. The outcome is a validated and prioritized set of KQPs that can serve as a decision-support tool for planners and policymakers seeking to improve the quality and integration of urban mobility systems.

The approach developed in this research builds on and extends previous applications of quality and sustainability assessment tools in transport studies. Among these, some have focused on the assessment of sustainable transport (Castillo & Pitfield, 2010), the quality of public transport services (Barabino et al., 2020), or micromobility systems (Carrara et al., 2021). In this research, the approach is applied for the first time to the multimodal dimension of urban mobility. This contributes to bridging the methodological gap between mode-specific evaluations and integrated system-level assessment.

The identification of KQPs thus represents a pivotal step in this research, as it provides the conceptual and operational foundation for the subsequent steps, which apply these parameters to empirical case studies for measuring and comparing the quality of urban mobility systems.

As illustrated in *Figure 20*, this phase of the research includes three main steps to select a narrow but representative set of KQPs.

The first phase is the "KQPs identification", which identifies a long list of potential KQPs from the previous literature review, defining evaluation components and related attributes for their assessment. The second phase is the "KQPs judgment", which involves academic experts to elicit their judgments about components, attributes, and KQPs. The third phase is the "KQPs assessment" and processes data to select the most promising set of KQPs. Sections 4.1.1, 4.1.2, and 4.1.3 describe each phase in detail.

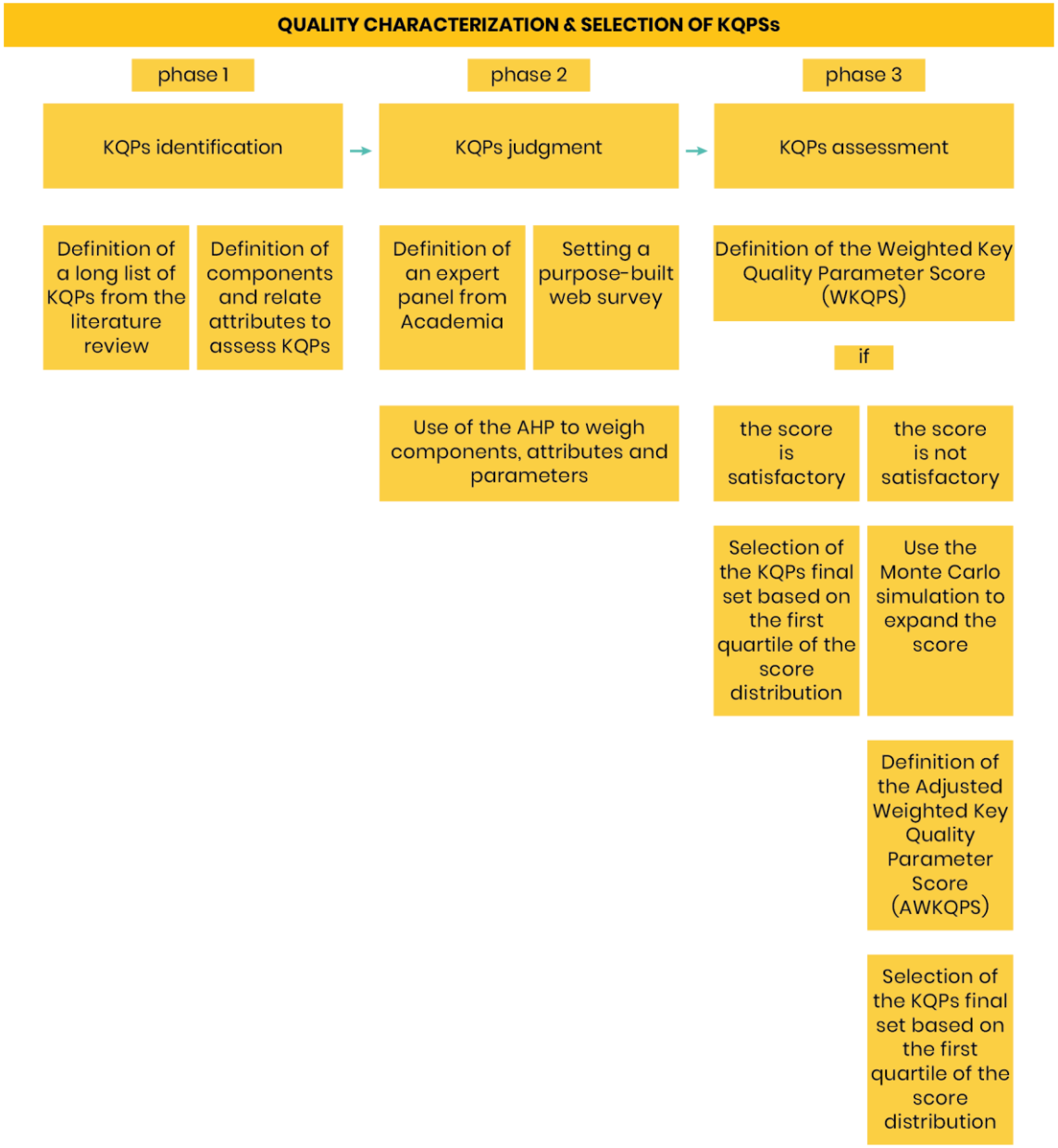


Figure 20. Flow chart for the selection of the representative set of KQPs.

4.1.1 Identification of KQPs

Phase 1 aims to identify the most effective parameters for assessing the quality of a multimodal and sustainable urban transport system. This phase includes two steps.

The first step identifies a long list of KQPs from the previous literature review. To structure it and ensure that all the most relevant quality aspects are included, the long list of KQPs is organised using the service quality norms EN 13816:2002 (CEN/TC 320, 2002) and EN 15140:2006 (CEN/TC 320, 2006) as a prior reference. Although these standards were originally developed to monitor the quality of service provided by Public Transport Companies (PTCs), adopting a structured classification based on recognized norms guarantees methodological consistency, transparency, and possible comparability of results. In this way, the criteria defined in these norms are adapted to be applicable to multimodal systems (*Table 6*). Based on this framework, each parameter of the long list will be clustered into eight macro-criteria. However, their relevance varies across transport systems, and their effectiveness in capturing the multimodal dimension is therefore uneven. In the best cases, a parameter applies to all, or at least three out of four transport systems considered; in the worst cases, it covers only one or two.

Table 6. Description of quality criteria.

Criteria	**Description**
Accessibility	It refers to the access to the transport system that includes interface with other transport modes.
Availability	It refers to the extent of the existing transport system considering the geography, time, frequency and transport modes.
Comfort	It refers to all the elements that make journeys relaxing and pleasurable.
Customer care	It refers to the service element introduced to affect the closest practicable match between the standard service and the requirements of any individual customer.
Environmental impact	It refers to the effect/impact on the environment resulting from a transport system.
Information	It refers to the systematic provision of knowledge about a transport system to assist in the planning and execution of the journey.
Safety & Security	It refers to the sense of personal protection experienced by travellers.
Time	It refers to all the aspects of time relevant to the journey's planning and execution.

The second step defines components and attributes to evaluate the relevance of quality parameters (*Table 7*). It considered the methodological quality of KQPs and their relevance in comparison to multimodal transport quality as components. The choice of these two components stems from the dual nature of this study’s challenge in selecting effective quality parameters. On one hand, parameters must be technically sound and usable within planning and decision-support processes, which requires clarity, measurability, and reliability; hence, the focus on methodological quality. On the other hand, since the objective of this research is to define and assess the quality of multimodal transport systems, it is essential that each parameter meaningfully captures the aspects that shape users’ experiences and influence urban mobility dynamics. For this reason, we included quality relevance as a specific evaluation component. Together, these components

ensure a balanced evaluation of each KQP, capturing both its operational feasibility and its significance in representing the quality of a multimodal transport system.

These components are aimed at identifying the main characteristics of KQPs, but they are not specific enough to drive the selection of KQPs. For this reason, they were both structured into different and more accurate attributes. Each attribute may have a specific level of importance that may vary to reflect different perspectives, although the components and related attributes are general and arise from KQPs considered as SMART, i.e., Specific, Measurable, Achievable, Realistic, and Timely (Doran, 1981).

As for the methodological quality, as suggested by Castillo and Pitfield (2010), we include the attributes of measurability, interpretability, speed, and easy availability. However, we add a fifth attribute that is critical for this research. It concerns the relevance of parameters to the assessment of a multimodal transport system. This attribute is crucial because the quality is evaluated across the entire travel experience, rather than just focusing on the journey with a specific service. This highlights the importance of considering every stage of the travel experience, from gathering pre-trip information to the final leg of the journey to the destination (Diana et al., 2016).

As for quality relevance, as suggested by Barabino and Di Francesco (2016), we include the attributes of integration between stakeholder and user perspectives, user orientation, measurability from both subjective and objective perspectives, and quantification of the number of users. Also in this case, we add a fifth attribute that is crucial to converge transport planning and urban planning, which concerns the relation between transport and public space quality. Indeed, the concept of quality cannot be limited to the transport mode but must be related to everything that interacts with it. Indeed, aspects of urban space can have a strong weight in influencing the travel mode choices of users. The absence of lighting, for example, can compromise users' perception of the security of a pedestrian or cycle route. Similarly, the presence of greenery or attractiveness can increase the perceived comfort of the route, favouring its choice.

Table 7. Components and attributes for assessing KQPs, adapted from Castillo and Pitfield (2010), Barabino et al. (2020), Carrara et al. (2021), and Carra et al. (2022).

Components	Attributes	Description
Methodological quality	Measurability	The parameters should be measured in a theoretically valid and reliable way.
	Ease of availability	It should be possible to collect data at a reasonable cost in line with the requirements of parsimonious Public Transport Companies (PTCs) and Public Administrations (PAs).
	Speed of availability	It should be possible to update the assessment regularly to reduce the time between consecutive measurements.
	Interpretability	It should be possible to consider unambiguous and universally understandable parameters. Parameters expressed in numerical values such as percentages or letters are preferred, being more user-friendly for technicians, senior managers and users.
	Relevance to multiple transport systems	The parameter should enable multiple transport systems to be simultaneously evaluated, rather than being measurable only for one system. The measurability of a parameter concerning multiple systems increases its effectiveness from a multimodal perspective.
Quality relevance	Integration between stakeholder and user perspectives	The evaluation of the quality using the parameter should re-orient the organisation as more user-focused. Stakeholders who provide and manage transport services and infrastructure (LPT provider - LPTP, sharing system providers -SSP, and public administration - PA) should consider the user as the key component, prioritizing their point of view.
	User-oriented	This is a challenging aspect of overall quality assessment, as when there are several parameters of a predefined quality aspect, it is important to avoid LPT/SSP/PA-oriented parameters that cannot be understood by users. For instance, in LPT, the measure of safety in terms of the “presence of usable handrail/handhold” versus “injured/km”.
	Measurability from subjective and objective perspectives	Both subjective and objective data should contribute equally to the quality monitoring process to ensure uniformity. The integration of subjective and objective parameters forces stakeholders - who provide and manage transport services and infrastructure - to assess the impact of performance on user satisfaction for each parameter. It also allows for collecting evidence on passenger perceptions of the services/infrastructure provided.
	Quantification of the number of users	The parameter should include a measure accounting for the number of users. The relevance of this attribute refers to EN 13816:2002 and EN 15140: 2006. Following these norms, the level of achievement of the objective should be expressed, where possible, as a ratio of transport users affected (CEN/TC 320 2002, 2006).
	Relation between transport and public space quality	The parameter allows us to measure aspects of both the quality of an urban transport system and the quality of public space. In this way, it can simultaneously provide assessments to support land use management and urban regeneration processes and improve the transport offer. For example, the presence of environmental facilities such as green spaces (parks, tree-lined streets, etc.) can be an aspect of transport quality that can promote active mobility; at the same time, it can be considered an aspect of public space quality, able to improve the perception of urban space and the management of land use and related environmental impacts.

4.1.2 Expert-based evaluation of KQPs

Phase two involves academic experts driving the selection of KQPs. The decision to involve only academic experts, rather than administrators or other stakeholders, is driven by the specific objectives of this phase. Academic experts provide a research-based and multidisciplinary perspective grounded in theoretical frameworks and comparative international experiences. Their role is to critically evaluate the conceptual validity and methodological robustness of the quality parameters. While administrators offer valuable operational and contextual knowledge, their views may vary significantly depending on local policies and constraints, which could introduce biases not aligned with the aim of

defining a universally applicable set of KQPs. Future research phases may involve other stakeholder groups to complement and validate these findings in practical contexts.

Academic experts are asked to assess components, attributes, and parameters through a purpose-built web survey that is structured using the PHP language. There are many methods for engaging academic experts, such as Contingent Field Survey (Blečić et al., 2016), Field and face-to-face Survey (Argin and Ozbil, 2015), Random Intercept Surveys (Rodriguez-Valencia et al., 2020). However, no single approach is universally applicable or accepted. In this case, a web survey is chosen because it allows for reaching many international academic experts at low operating costs, such as interview time, data coding, and data processing. Additionally, a web survey is non-intrusive and requires less effort from respondents.

Parallelly, measuring the relevance of components and related attributes is a complex task that requires an in-depth knowledge of the subject. Lacking technical knowledge, transport users and travellers were not directly considered here, even though their perspective is a priority in this research. Their perspective will become crucial in assessing both the expected and perceived quality of the parameters selected using the proposed approach.

Involved academic experts emerged from academia and literature review and include university professors and researchers working on the topics of mobility and quality assessment of urban transport systems. Because the research investigates a multimodal dimension, academic experts in different fields of transportation were involved: pedestrian mobility, cycling, public transport, and micromobility.

However, this heterogeneity may lead to evaluations conditioned by specific cultural and technical backgrounds that are not always multimodality-oriented, potentially resulting in biased assessments or inconsistent values. To objectively assess the relevance of components and related attributes, we apply a weighting process, thereby minimising the likelihood that the number of academic expert judgments would also increase the possibility of receiving biased opinions. According to Zelany (1974) and Castillo and Pitfield (2010), weights can be directly assigned by asking experts to express their preferences for individual items. However, this approach may be problematic, as humans often struggle to translate complex information into consistent weights when faced with numerous items. Moreover, other approaches could be adopted (e.g., Wang and Lee, 2009). Nevertheless, our approach includes the AHP, which mitigates the risks of bias (Saaty, 1977, 1987, 1994; Wind & Saaty, 1980), establishing weights from a pairwise comparison of items and generating a ratio scale from each set of comparisons to deal with inconsistency in judgments (Saaty, 1977, 1987, 1990, 1994)[1].

[1] Although the method has been subject to criticism (Dyer, 1990), numerous scholars have supported its application and highlighted the benefits associated with its use (e.g., Harker and Vargas, 1990; Saaty, 1990; Forman and Gass, 2001; Ramanathan, 2001; Millet and Wedley, 2002; Macharis et al., 2004; Oguztimur, 2011). Furthermore, the Analytic Hierarchy Process (AHP) is widely employed across various engineering disciplines, which has further encouraged its adoption among bus operators (e.g., de Steiguer et al., 2003).

In this process, the AHP method utilized subjective pairwise comparisons of quality components or attributes, which were then aggregated into weights, accounting for the varying degrees of subjectivity among the academic experts involved. The approach was particularly effective in resolving disagreements between academic experts with potentially differing viewpoints. A pairwise comparison matrix was created for each academic expert, where the rows and columns represented the quality components or attributes, and each entry reflected the weight of one component/attribute relative to another.

Subsequently, a weight vector was calculated and normalized for each quality component or attribute. Any inconsistencies in the judgments were then identified, prompting a consistency test to ensure the reliability of the assessments within each matrix.

To be more formal, let:

- S be the set of components or attributes;
- J be the set of academic experts involved;
- $v_{sq,j}$ be the numerical judgment provided by an academic expert j ∈ J of the pairwise comparison between a pair of components or attributes s ∈ S and q ∈ S;
- V_{sj} be the overall un-normalised weight of component or attribute s ∈ S returned by academic expert j ∈ J;
- $CI_{(j)}$ be the consistency index and assesses whether the judgments provided by academic expert j ∈ J are consistent with the choices reported in the survey;
- $\lambda_{\max(j)}$ be the maximum eigenvalue required for computing the measure of consistency by academic expert j ∈ J;
- RCI be the random consistency index, which represents a CI function tabulated based on the maximum number of quality components or attributes (Saaty, 1980).

A four-step algorithm was used to calculate the weights and conduct a consistency check on the judgments.

For every academic expert j ∈ J:

1. Construct the matrix of pairwise comparison V_j between components or attributes for each academic expert j ∈ J;
2. Compute V_s and v_s from this matrix. Specifically, the computation of the weight vector V_s is performed as follows (among the different methods):

$$V_{sj} = \sqrt[|S|]{\prod_{s,q \in S} \frac{v_{sq,j}}{v_{qs,j}}} \qquad \forall j \in J \tag{1}$$

Next, V_s is normalised using the average arithmetic method as follows:

$$v_{sj} = \frac{V_{sj}}{\sum_{s \in S} \quad sj} \qquad \forall j \in J \tag{2}$$

3. Check the consistency, computing $\lambda_{\max(j)}$ as follows:

$$\lambda_{max(j)} = \frac{\sum_{q \in S} \left[\frac{\sum_{s \in S} \left(\frac{v_{sqj}}{v_{qsj}} * v_{sqj} \right)}{v_{qsj}} \right]}{|S|} \quad \forall j \in J \tag{3}$$

4. Check if $\lambda_{\max(j)} \geq |S|$ and compute CI_j as follows:

$$CI_j = \frac{(\lambda_{max(j)} - |S|)}{(|S|-1)} \qquad \forall j \in J \tag{4}$$

where the assessments exhibit perfect consistency if $\lambda_{max(j)} = |S|$, thus $CI_j = 0$.

Finally, compute RCI as follows:

$$RCI_j = \frac{CI_j}{RCI} \qquad \forall j \in J \tag{5}$$

Pairwise comparisons are considered trustworthy when the CRI ≤ is 0.1 or when expert judgments demonstrate coherence and align with the survey responses. If the CRI is > 0.1, experts are asked to revise their evaluations due to the inconsistency in their judgments. Detailed guidance on the application of the Analytic Hierarchy Process (AHP) is provided in Saaty's publications (Saaty, 1987; 1994). For each expert j ∈ J, separate matrices are developed at the component or attribute levels. Experts whose responses fail to meet the required consistency standard are excluded from the analysis—only consistent judgments are used in the weighting process. The final weights are calculated by averaging the weights derived from all consistent expert inputs. These last weights are established for each component and attribute and serve as inputs for the following step.

In this study, the AHP is applied twice: first, academic experts compare the two quality components, and, next, the two sets of five attributes characterizing each component. They adopt a numerical scale from 1 to 9.

Next, academic experts rate each KQP against all the attributes according to predefined scales. In this case, we use a shorter numerical scale from 1 to 5. Compared to the first phase, the second one took longer for academic experts to assign values to the long list of parameters derived from the literature across the 10 quality attributes. So as not to discourage academic experts from contributing to the survey and trying to get more responses, the range was reduced from 1-9 to 1-5. The narrower range in the second stage of the survey made it easier for academic experts to evaluate, as it sped up the judgment process by reducing the number of nuances to consider.

4.1.3 Assessment framework for KQPs

This stage processes the data collected among the academic experts and returns the best score associated with the KQPs. This score is defined as the Weighted Key Quality Parameter Score (WKQPS) and is obtained from an additive weighting, which aggregates weights and outcome marks to evaluate the performance of each KQP. Specific steps characterise this stage, as shown below.

To be more formal, let:

- P be the set of KQPs;
- *H* be the set of the methodological attributes;
- *K* be the set of the quality attributes;
- m_j be the weight of the methodological quality component as derived by AHP, according to the evaluation of academic expert j ∈ J;
- q_j be the weight of the quality relevance component as derived by AHP, according to the evaluation of academic expert j ∈ J;
- w_{jh} be the weight of attribute h ∈ H as derived by AHP, according to the evaluation of academic expert j ∈ J;
- w_{jk} be the weight of attribute k ∈ K as derived by AHP, according to the evaluation of academic expert j ∈ J;
- $\underline{M}$ be the average weight of the methodological quality component;
- $\underline{Q}$ be the average weight of the quality relevance component;
- $\underline{w_h}$ be the average weight of attribute h ∈ H;
- $\underline{w_k}$ be the average weight of attribute k ∈ K;
- V_{pjh} be the mark of parameter p ∈ P for attribute h ∈ H according to the evaluation of academic expert j ∈ J;
- V_{pjk} be the mark of parameter p ∈ P for attribute k ∈ K according to the evaluation of academic expert j ∈ J;
- $\underline{V_{ph}}$ be the average mark of parameter p ∈ P for attribute h ∈ H;
- $\underline{V_{pk}}$ be the average mark of parameter p ∈ P for attribute k ∈ K.

For each parameter *p*, the WKQPS is computed according to the following four-step algorithm.

1. Compute $\underline{M}$ and $\underline{Q}$

$$\underline{M} = \frac{\sum_{j \in J} m_j}{|J|} \tag{6}$$

$$\underline{Q} = \frac{\sum_{j \in J} q_j}{|J|} \tag{7}$$

2. Compute $\underline{w_h}$ and $\underline{w_k}$

$$\underline{w_h} = \frac{\sum_{j \in J} w_{jh}}{|J|} \forall h \in H \quad (8)$$

$$\underline{w_k} = \frac{\sum_{j \in J} w_{jk}}{|J|} \forall k \in K \quad (9)$$

3.

4. Compute $\underline{V_{ph}}$ and $\underline{V_{pk}}$

$$\underline{V_{ph}} = \sum_{j \in J} \frac{V_{pjh}}{|J|} \forall h \in H;\ \forall p \in P \quad (10)$$

$$\underline{V_{pk}} = \sum_{j \in J} \frac{V_{pjk}}{|J|} \forall k \in K;\ \forall p \in P \quad (11)$$

5. Compute $WKQPS_p$

$$WKQPS_p = \underline{M} \cdot \left(\sum\nolimits_{h \in H} \underline{w_h} \cdot \underline{V_{ph}}\right) + \underline{Q} \cdot \left(\sum\nolimits_{k \in K} \underline{w_k} \cdot \underline{V_{pk}}\right) \forall p \in P \quad (12)$$

Lastly, we determined the final set of KQPs by sorting the list in descending order of WKPQS, where the best parameters correspond to their highest estimated values.

First, the first quartile of the distribution of WKPQS values of the overall set of parameters is chosen, considering the highest-ranked KQPs relevant for the quality management of multimodal and sustainable transport systems in urban areas. The first quartile is a reasonable way to provide an adequate assessment of the quality of an urban transport system because:

- it is not excessively high, so it does not require excessive efforts in terms of time and resources;
- it is not excessively low, so it is not in danger of being unrepresentative and does not oversimplify a very complex system.

Second, the parameters higher up the list were supplemented and modified to meet the following selection criteria:

- at least one parameter for each of the 8 criteria derived from standards EN 13816:2002 (CEN/TC 320, 2002) and EN 15140:2006 (CEN/TC 320, 2006) must be present in the final set;
- each parameter must be able to measure at least three of the four transport categories included in the study (public transport, pedestrian mobility, bicycle mobility, and electric micromobility), in the final set. This criterion was chosen to obtain a set that is as applicable as possible to a multimodal transport system.

Finally, the proposed methodology includes a final uncertainty condition to address potential limitations in the sample size obtained from the web-based survey (if any).

Specifically, if the number of expert responses is sufficiently large and balanced, the final set of KQPs is selected based on the first quartile of the WKQPS distribution. However, when the number of responses is relatively low, an adjustment of the values of the WKQPS is introduced to ensure the robustness of the results and to mitigate coverage bias.

The adjustment of values of the WKQPS may become necessary because web surveys, while efficient and scalable, are particularly prone to bias when the respondent pool is limited in size or composition (Dever et al., 2008). This bias can distort the aggregated evaluations and affect the reliability of parameter ranking. To overcome this limitation, we apply Monte Carlo simulation, as previously done in similar works on transit service quality (Barabino et al., 2020) and electric charging infrastructure (Carra et al., 2022) and other (Castillo and Pitfield, 2010). In this context, the deterministic evaluation model, based on actual expert scores, is converted into a stochastic model. A set of random numbers is generated and used as input in an iterative process that simulates the full computation of WKQPS, producing an Adjusted WKQPS (AWKQPS). This procedure is repeated several times to generate a simulated data distribution that reflects the same structure as the expert responses. A probability distribution is chosen to represent the available knowledge, and the final ranking is obtained by reapplying the evaluation model to the simulated data. Through this step, the methodology not only increases the robustness of the final selection but also enhances its reliability in scenarios where expert participation is limited, ensuring that the selected KQPs maintain statistical validity and conceptual soundness.

Finally, to further investigate the relevance of the best KQPs, we compared the results with ranking methods, evaluating whether the best-ranked KQPs are considered equally important in the surveyed literature conducted with the PRISMA statement. The former ranking method consists of a Bibliometric Score, i.e., the number of times the paper that used each indicator has been cited. The latter ranking method uses a Citation Score, i.e., a count of how many times each parameter appears in the literature analysed (i.e., at most once per paper).

4.2 Measurement of quality in multimodal transport systems

The second phase of the methodology concerns the empirical measurement of quality. The measurement of quality in multimodal transport systems represents a fundamental stage within the broader quality-monitoring and management cycle. A comprehensive evaluation must capture not only how services perform in practice, but also how they are conceived, expected, and experienced. This logic is formalised in the quality loop, originally introduced within the European Fourth RTD Framework Programme (QUATTRO, 1998) and subsequently codified in the European Standard EN 13816 (CEN, 2002).

The quality loop conceptualises service quality through four interdependent dimensions:

- **Designed quality**: the level of service established by responsible organisations (e.g., public transport operators, shared micromobility providers, public authorities) through strategic and operational planning;
- **Expected quality**: the level of service that users expect and consider satisfactory;
- **Delivered quality**: the actual level of service provided and observed in real operating conditions;
- **Perceived quality**: users' subjective assessment of the service based on their direct experience.

These four dimensions are mutually reinforcing and should be interpreted from a user-centred perspective, focusing on mobility needs and expectations rather than on organisational priorities alone. In multimodal urban systems, where multiple actors jointly shape mobility provision, the integrated consideration of these four dimensions becomes essential to understanding quality gaps, enhancing coordination, and supporting evidence-based improvements.

Consistent with established practice in transport quality research, this study adopts an integrated approach combining objective and subjective quality assessments. Relying solely on operational indicators risks overlooking the experiential dimension of mobility, whereas focusing only on perception may obscure structural limitations. Instead, quality emerges from the interaction between system performance and user experience.

Two complementary measurement techniques are therefore employed.

The first is the Mystery Client (MC) technique, which provides an objective evaluation of delivered quality by documenting service conditions as they are experienced in real operating environments. Trained evaluators follow predefined observation protocols and measure all the predefined KQPs, offering an evidence-based perspective on actual system quality from the user's standpoint.

The second technique is the Customer Satisfaction (CS), which captures subjective quality by gathering users' judgements on both expected and experienced service levels. Respondents rate the importance of each quality parameter and their satisfaction, enabling the identification of gaps between what users anticipate and what they perceive during everyday mobility. The survey design draws on established transport-quality and customer-experience research.

Integrating the insights obtained through MC observations with CS responses allows a comprehensive assessment of multimodal transport systems. This dual perspective makes it possible to identify areas where objective performance and user perception align, as well as situations where discrepancies emerge. Such convergence and divergence patterns provide valuable evidence for uncovering latent quality issues, understanding behavioural responses, and prioritising targeted interventions to improve multimodal mobility. *Figure 21* summarizes the workflow of the quality measurement phase of the overall methodology.

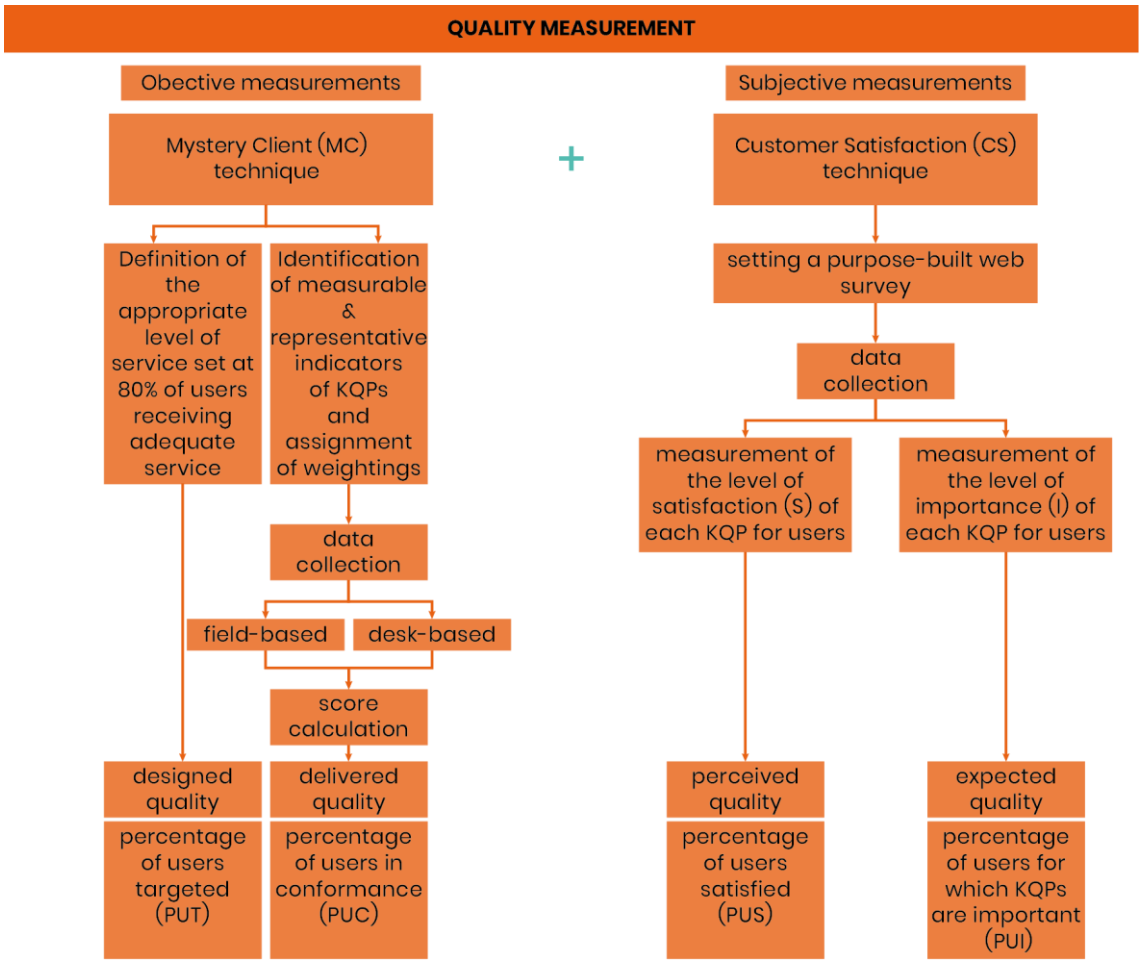


Figure 21. low chart for the quality measurement including objective and subjective perspectives.

4.2.1 Objective measurements

The objective measurements of quality were conducted using the Mystery Client technique, which constitutes the cornerstone of the objective assessment phase, enabling the systematic evaluation of service and infrastructure quality under real operating conditions. This method allows the researcher to observe the transport system in a user-centred manner. Following the identification and prioritisation of KQPs, the approach required the translation of these parameters into measurable indicators to support structured field observation. A set of observable and quantifiable indicators was therefore defined for each KQP. The selection process was guided by two complementary principles:

- alignment with constructs and operational practices recurrent in the transport-quality literature;
- feasibility of systematic in-field measurement across different urban contexts and mobility modes.

This ensured that the chosen indicators were both theoretically grounded and operationally viable, enabling systematic application within the Mystery Client framework.

Priority was given to indicators capable of capturing operational, infrastructural, and experiential dimensions in a clear and replicable manner.

Each indicator was operationalised through the assignment of a measurement scale, scoring rules, and a weighting coefficient, enabling standardised evaluation across transport modes and case studies. Weighting reflected the relative importance of each indicator within the corresponding quality criterion.

Designed quality

Designed quality represents the percentage of users who are expected to receive the desired level of service. We denote this target as percentage of users targeted (PUT). It is important to define a minimum PUT percentage to guarantee a basic level of service quality. This threshold can be adjusted according to the relative importance of specific quality parameters.

In defining the designed quality, a baseline quality objective was established, specifying that at least 80% of users should receive a level of service considered adequate. This threshold represents the minimum acceptable standard of service provision and serves as a reference for subsequent assessment and management phases. The 80% benchmark is consistent with values commonly adopted in transport-quality certification frameworks, such as those defined by AFNOR Certification for urban public transport services (AFNOR, 2005).

Delivered quality

Delivered quality refers to the objective measurement of the service as it is actually provided under real operating conditions. This dimension is assessed across all previously selected KQPs and, crucially, from the user's perspective. This approach is essential for two main reasons. First, the parameters under examination must be meaningful and interpretable by users. Second, quality must be measured in a way that reflects how users effectively experience and receive the service.

To capture delivered quality, this study employs the Mystery Client (MC) technique[2], a structured observational method in which trained evaluators travel through the system as ordinary users while systematically assessing predefined quality parameters using a standardised checklist (Appendix A, *Figure A.1*). Each KQP is operationalised through one or more measurable indicators, for which a scoring rule, a weight, and a maximum attainable weight are defined. Indicators are evaluated using a binary classification (good / bad quality), thereby facilitating rapid in-field assessment and efficient data processing.

The weighting methodology may vary depending on the objectives and stage of development of the quality-assessment framework. In an initial exploratory phase,

[2] All Mystery Client measurements were performed by the author as part of the doctoral research activity.

indicator weights can be assigned through a structured, expert-informed judgement process aimed at ensuring internal coherence among criteria and indicators and at testing the operational feasibility of the assessment protocol.

In more advanced applications, a more formalised multi-criteria procedure may be adopted. Pairwise comparisons among indicators within each criterion can be conducted following principles inspired by the Analytic Hierarchy Process (AHP) or similar frameworks. Weights derived from expert evaluations may then be aggregated - typically through arithmetic averaging - and their stability assessed using measures such as standard deviation or coefficient of variation. Low dispersion values indicate a consistent and reliable weighting structure.

In applications where delivered quality is assessed exclusively through MC observations and no user-level sample is available, the computation of PUC requires an operational adaptation. Since the MC acts as a single informed observer rather than representing a population of users, the Percentage of Users in Conformity is computed as a proxy measure. This proxy expresses the observed quality as a normalised percentage of the maximum attainable score for each KQP. In this case, the delivered quality percentage (PUC_MC) is obtained by dividing the MC-observed score by the corresponding maximum score and multiplying the result by 100. This value, therefore, reflects the share of the designed quality standard that is effectively delivered on the ground, rather than the share of users for whom the standard is met. The theoretical definition of PUC, presented in Chapter 4, remains unchanged; what varies is the operational procedure used to estimate it in the absence of user-based data.

Once the field data have been collected, delivered quality is quantified through the Percentage of Users in Conformity (PUC). This metric expressing the share of users who receive a level of service that meets the predefined quality standard established with the designed quality.

When O-D connections cover relatively large areas or extend over long distances, assessing quality indicators along the entire route as a single unit may lead to distorted or uninformative results. For example, binary or low-resolution measurement scales may overestimate the presence or relevance of certain features when they occur only sporadically along an extended corridor. To ensure representativeness and maintain the granularity of the assessment, routes can be subdivided into smaller segments of manageable and consistent length. This segmentation - typically performed using GIS tools - allows each portion of the route to be evaluated independently, capturing variations in infrastructure, services, and environmental conditions along the corridor.

Building on this structure, delivered quality is then computed through a hierarchical process that mirrors how users experience mobility in real conditions. Measurements are first obtained at the segment level (the smallest travel unit), then aggregated across paths (alternative sequences of segments connecting the same origin-destination pair), and finally synthesised at the O-D level. Since each criterion comprises multiple KQPs, a final

step aggregates KQP-level scores into an overall criterion-level delivered quality measure. This multi-layered approach ensures that local conditions are accurately reflected while still producing meaningful quality assessments at higher spatial and functional scales.

When the PUC proxy (PUC_MC) is used, the hierarchical aggregation procedure remains unchanged: segment-level scores are normalised against their maximum values, then aggregated to paths, O-D pairs, and finally criteria using the established weighting structure. The result is a delivered-quality percentage that maintains interpretability and internal consistency while acknowledging the observational nature of the data source.

Figure 22 schematically illustrates the multilayer structure.

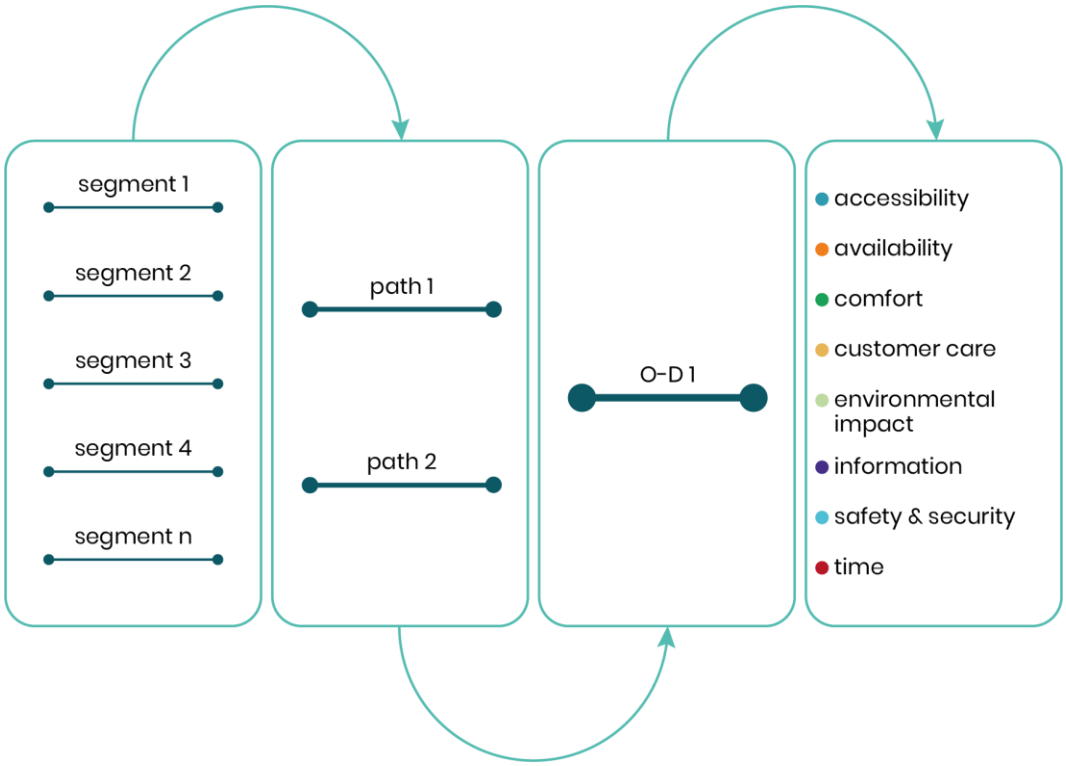


Figure 22. Hierarchical aggregation process for delivered quality: from segment-level observations to path-level, O-D-level, and criterion-level scores.

To be more formal, let:

- c be the quality criterion (e.g., Accessibility, Comfort, Safety & Security);
- $P(c)$ be the set of KQPs belonging to criterion c;
- $I(p)$ be the set of indicators belonging to KQP p;
- $T(o,d)$ be the set of path belonging to O-D pair o,d;
- $L(t)$ be the set of segments belonging to path $t \in T(o,d)$;

- τ be the conformity threshold corresponding to the designed quality objective (e.g.,0.8; 800; 80%);
- x_p^l be the binary indicator of segment l for KQP p (=1 if the segment $l \in L(t)$ satisfies the quality standard for KQP p; 0 otherwise);
- U_l be the number of users travelling on segment $l \in L(t)$;
- W_p be the weight of KQP p within criterion c;
- W_i be the weight of indicator i within KQP p;
- obs_l be the MC observation associated with the segment$l \in L(t)$.

The PUC at the level of segment, path and OD, is computed as follows:

1. Compute the $Rate$ of the observation obs_l of the segment l for each KQP p

$$Rate_p^{obs} = \sum_{i \in I(p)} W_i \tag{13}$$

2. Compliance check (segment level):

$$x_p^{obs} = \begin{cases} 1 \; if \; Rate_p^{obs} \geq \tau \\ 0 \; otherwise \end{cases} \qquad \forall p \in P_{(c)} \; , \; \forall l \in L_{(t)} \tag{14}$$

3. Segment-level delivered quality

$$PUC_p^l = \frac{\sum_{obs \in OBS(l)} x_p^{obs} * U_p^{obs}}{\sum_{obs \in OBS(l)} U_p^{obs}}; \forall p \in P_{(c)} \; , \; \forall l \in L_{(t)}, \forall t \in T_{(o,d)} \tag{15}$$

4. Path-level delivered quality (PUC_p^t)

   For each KQP p, the Percentage of Users in Conformity on segment s is computed as:

$$PUC_p^t = \frac{\sum_{l \in L(t)} PUC_p^l}{|L(t)|}; \forall p \in P_{(c)} \; , \; \forall l \in L_{(t)}, \forall t \in T_{(o,d)} \tag{16}$$

5. O-D-level delivered quality (PUC_p^t)

   A path typically consists of multiple segments, each potentially characterised by different quality levels. The average delivered quality for KQP p on path t is therefore:

$$\overline{PUC_p^t} = \frac{1}{|T_{(od)}|} \sum_{t \in T(o,d)} PUC_p^t \qquad \forall p \in P_{(c)} \; , \; \forall t \in T_{(od)} \tag{17}$$

Building on this structure, delivered quality is computed through a set of hierarchical formulas. At the segment level, each KQP is assessed using a binary indicator that identifies whether the segment satisfies the predefined quality threshold. These values are then averaged across all segments belonging to a given path, producing a path-level delivered quality score. When more paths connect the same origin-destination pair, their results are aggregated to obtain an O-D-level measure. Since each quality criterion comprises multiple KQPs, a final weighted aggregation yields the delivered quality for each criterion. The outcome of this process is the Percentage of Users in Conformity (PUC), which quantifies the share of users who receive a service that meets the designed quality standard across segments, paths, O-D pairs, and criteria. The full set of equations is reported above, while *Figure 22* illustrates the hierarchical structure supporting this computation.

4.2.1.1 Preliminary Assessment Through a Simplified Delivered Quality Procedure

Before implementing the full delivered quality methodology - based on binary conformity assessment, user-based weighting, and hierarchical aggregation from segments to paths and origin-destination pairs - a simplified preliminary version can be employed. This initial step serves two main purposes:

- to validate the operational feasibility of the Mystery Client protocol;
- to assess the clarity, observability, and discriminative capacity of the selected indicators in real conditions.

In this simplified variant, delivered quality is measured directly at the level of individual routes or segments, without reconstructing full multimodal journeys or aggregating across the hierarchical structure. Instead of computing the Percentage of Users in Conformity (PUC), quality is represented through a weighted multi-criteria score, where each indicator contributes proportionally to its assigned weight.

To be more formal, let:

- c be the quality criterion (e.g., Accessibility, Comfort, Safety);
- $P\ (c)$ be the set of KQP belonging to criterion c;
- $I(p)$ be set of the indicators belonging to KQP p;
- S_{pi} be the score observed for indicator i belonging to KQP p;
- W_{pi} be the corresponding weight.

The simplified score for each assessed route is computed as:

$$S_{route} = \sum_{c} \sum_{p \in P(c)} \sum_{i \in I(p)} S_{pi} W_{pi} \quad (18)$$

In this application, indicator scores were predominantly expressed using simplified scales (mostly binary, e.g., 0/10, or ternary, e.g., 0/5/10) to facilitate rapid and reliable in-field observation. Weights were assigned by the author on the basis of literature review and operational judgement rather than through a formal multi-criteria elicitation process. This allowed for the immediate deployment of the indicator system and served as a testing ground for its operational robustness. This preliminary assessment does not incorporate user-based weighting or hierarchical aggregation across segments, paths, and O-D pairs.

Nevertheless, it plays a crucial methodological role. It enables the refinement of indicators, scoring rules, and measurement procedures prior to implementing the full delivered-quality framework, where the computation of the Percentage of Users in Conformity (PUC) becomes central. Once this initial validation phase is completed, the comprehensive methodology can be applied to assess multimodal journeys in a manner that is conceptually rigorous, scalable, and consistent with user-oriented quality standards.

4.2.2 Subjective measurements

The subjective measurements of quality were conducted using the Customer Satisfaction (CS) technique, a well-established approach in transport research for capturing users' expectations, experiential evaluations, and overall satisfaction levels (Parasuraman et al., 1988; Eboli & Mazzulla, 2009; de Oña & de Oña, 2014). Unlike objective assessments, which measure the service as it is delivered, CS methods focus on how the level of service is perceived and valued by users. This perspective is essential, as perceived quality strongly influences travel behaviour, acceptance of transport policies, and the willingness to adopt sustainable mobility options.

Within this methodological framework, respondents are asked to evaluate the set of KQPs. The CS approach distinguishes between two complementary dimensions: importance, which reflects the relevance that users attribute to each parameter, and satisfaction, which represents the extent to which these expectations are met by the transport system. Comparing importance and satisfaction enables the identification of quality gaps, highlighting those attributes where system quality falls short of user expectations and where targeted interventions may be most effective.

The Customer Satisfaction approach served three key purposes within the research:

- validation of the KQPs from a user-centred perspective, complementing the expert-driven prioritisation carried out in earlier phases;
- measurement of perceived and expected quality across mobility modes, allowing comparison with objective field observations from the Mystery Client technique;

- identification of intervention priorities, supporting the quality-management phase through the detection of gaps between expected and experienced quality.

By systematically integrating subjective perceptions with objective observations, the methodology enables a multidimensional and user-centred understanding of multimodal transport quality. This dual perspective is essential in contemporary mobility planning, where service design increasingly requires alignment between technical performance and the lived experience of diverse user groups.

Expected quality

Expected quality represents the level of importance attributed by users to each KQP. This is quantified through the Percentage of Users for whom the parameter is Important (PUI). After data collection, each respondent evaluates the relevance of every KQP using a predefined rating scale. These ratings are then converted into a binary variable indicating whether the parameter is considered important or not, based on a threshold established ex ante. Some parameters can be directly associated with quantifiable service features, while others relate to experiential or perceptual aspects. In these latter cases, user judgement becomes essential and must be translated into a systematic and comparable metric. The binary classification allows both types of aspects - whether more measurable or more perception-based - to be incorporated consistently into the analysis. It simplifies the evaluation process for respondents and ensures that the subsequent aggregation remains methodologically coherent across different categories of indicators.

To be more formal, let:

- U be the set of users;
- GI be the judgment of importance of the KQP p given by the $u \in U$ who moves using the using t $\in T(o,d)$;
- $GI_{t,p}^{(u)} = 1$ if user $u \in U$ moving using t $\in T(o,d)$ considers the KQP p important based on an importance threshold, and 0 otherwise;

For each path $t \in T(o,d)$ and each KQP p, the expected quality is computed as:

$$PUI_{t,p} = \frac{1}{|U|} \sum_{u \in U} GI_{t,p}^{(u)} \cdot 100 \quad (19)$$

Perceived quality

Perceived quality describes how users evaluate the service based on their direct experience. It represents the final component of the quality loop and captures the subjective assessment of the same KQPs considered in the previous phases. The measurement process follows an approach analogous to that used for expected quality, with the difference that users are asked to rate their satisfaction rather than the importance of each parameter.

Participants assign a satisfaction score using an ordinal Likert scale. To ensure consistency and comparability across parameters and case studies, the original ratings are converted into a binary variable based on a predefined threshold: scores at or above the threshold indicate satisfactory quality, whereas scores below it indicates dissatisfaction. This transformation simplifies the analysis and enables a direct interpretation of perceived quality as the share of users whose experience meets an acceptable standard.

For each path t and for each KQP p, perceived quality is computed as the Percentage of Users Satisfied (PUS).

To be more formal, let:

- U be the set of users;
- GS be the judgment of satisfaction of the KQP p given by the $u \in U$ who moves using the using t $\in T(o,d)$;
- $GS_{t,p}^{(u)} = 1$ if user $u \in U$ moving using t $\in T(o,d)$ assigns to the KQP p a satisfaction score greater than or equal to a predefined threshold, and 0 otherwise;

$$PUS_{t,p} = \frac{1}{|U|} \sum_{u \in U} GS_{t,p}^{(u)} \cdot 100 \tag{20}$$

Survey Design and Administration

In this research, expected and perceived quality were investigated through an online survey. It was developed and administered to residents of the city of Verona (Italy), with the objective of evaluating citizens' perceptions of the quality of the urban transport system.

Respondents were asked to assess their daily mobility experience, specifically focusing on trips between their residential neighbourhoods and destination neighbourhoods where workplaces, educational institutions, and recreational or sports facilities are

concentrated. The survey centred on the transport modes most relevant for sustainable urban mobility:

- non-motorised transport (walking and cycling)
- LPT (urban bus services)
- electric micromobility (shared or private e-scooters and e-bikes)

Participants evaluated the KQPs previously identified through the systematic literature review and expert-based weighting process. For each parameter, two separate ratings were collected on a 1-10 Likert scale:

- importance rating - measuring the perceived relevance of the parameter
- satisfaction rating - reflecting the respondent's experience

This dual-rating structure allowed a direct comparison between expected and perceived quality, supporting the identification of quality gaps and priority areas for policy and operational improvement.

The survey was anonymous and disseminated via institutional communication channels, community networks, and social media platforms. Data collection took place in June-July 2025, and responses were analysed in aggregate form. While not designed to be statistically representative of the entire population, the sampling strategy ensured the inclusion of a wide range of demographic groups and mobility profiles, consistent with the exploratory and methodological validation aims of the study.

4.3 Management of quality in multimodal transport systems

Quality management represents the final phase of the assessment framework and aims to identify critical areas while guiding both strategic and operational improvements. In line with established approaches in the transport-quality field (e.g., EN 13816, 2002; Barabino & Di Francesco, 2016), this phase analyses the relationships among the four components of the quality loop, with each component corresponding to a specific gap.

The first quality gap (QG1) concerns the difference between designed and expected quality and indicates the extent to which service providers define standards that truly reflect users' expectations. The second gap (QG2) relates to the difference between delivered and designed quality and serves to measure the ability of operators and public authorities to achieve the objectives they have set. The third gap (QG3) addresses the difference between perceived and delivered quality, highlighting whether objectively observed quality aligns with users' actual experience. The fourth gap (QG4) compares expected and perceived quality and provides insight into overall user satisfaction and the degree to which services meet mobility needs and expectations. *Figure 23* summarizes the workflow of the quality management phase of the overall methodology.

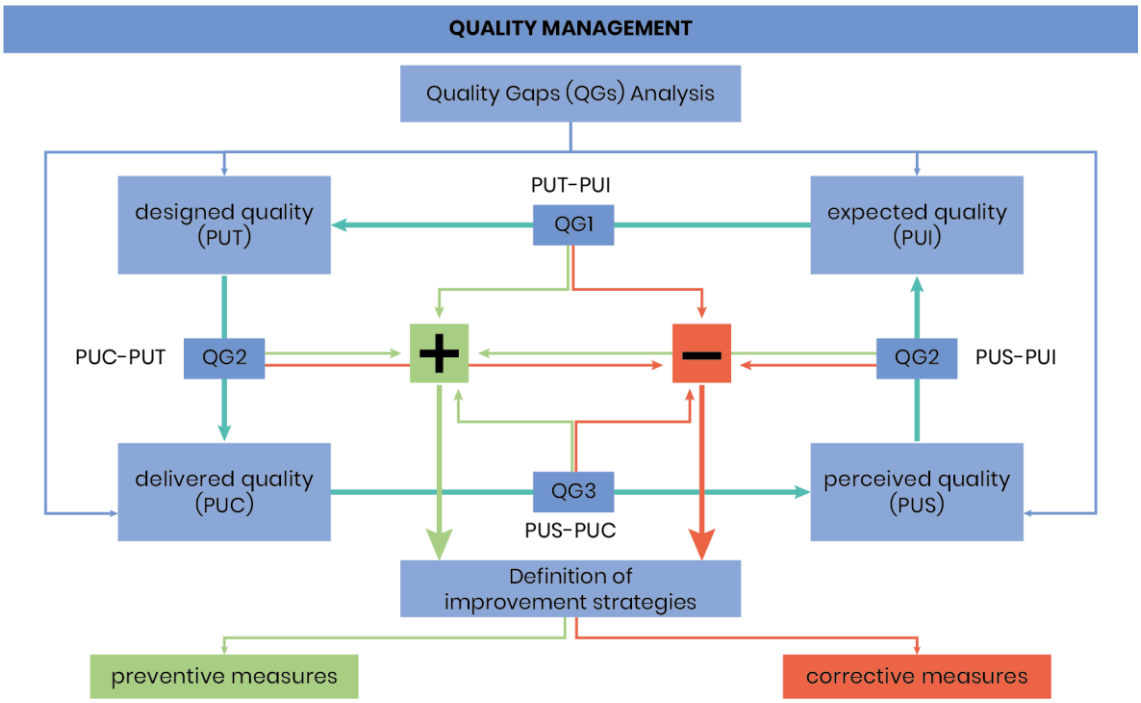


Figure 23. Flow chart for the quality management based on the quality gap analysis.

Through this perspective, gap analysis supports the identification of weaknesses and priority areas for improvement. While several studies have explored individual gaps - particularly QG4 in relation to user satisfaction and behavioural responses (e.g., de Oña & de Oña, 2014) - systematic examination of multiple gaps remains relatively limited. This gap in the literature reinforces the need for integrated approaches to quality management, especially in complex and multimodal mobility systems.

4.3.1 Quality Gaps Analysis

In this phase, the methodology quantifies the differences between each pair of quality dimensions, allowing the identification of critical mismatches and areas where service improvement is required. Each gap represents a specific misalignment in the quality loop and provides insight into where preventive or corrective measures should be prioritised.

Quality Gaps QG1 (PUT-PUI) and QG2 (PUC-PUT) reflect discrepancies within the service providers' or/and local authorities' sphere (including PTCs, shared micromobility operators, and PAs), at the strategic and operational levels, respectively. Quality Gaps QG3 (PUS-PUC) and QG4 (PUS-PUI), instead, capture misalignments between the service as delivered and users' perceptions or expectations.

To be more formal, let:

- PUT is the percentage of users targeted;
- PUI is the percentage of users for whom the parameter is important;

- PUC is the percentage of users in conformity;
- PUS is the percentage of users satisfied.

The four gaps are computed as follows:

$$QG1 = PUT - PUI \quad (21)$$

$$QG2 = PUC - PUT \quad (22)$$

$$QG3 = PUS - PUC \quad (23)$$

$$QG4 = PUS - PUI \quad (24)$$

QG1 captures the first misalignment between users and service providers or/and local authorities at the tactical level. It is calculated as the difference between the designed quality (PUT), set by service providers and local authorities, and the expected quality (PUI), expressed by users. A negative value signals that service standards have been defined below what users consider important, indicating a strategic mismatch in planning priorities.

QG2 measures the discrepancy between the quality standards set by service providers and local authorities (PUT) and the quality actually delivered (PUC). This gap reflects the service providers' and local authorities' ability to meet predefined targets. Larger negative values point to operational shortcomings, whereas positive values indicate that the service exceeds the established standards.

QG3 represents the second point of interaction between service providers and local authorities, and users, at the operational level. It compares perceived quality (PUS) with delivered quality (PUC), thus assessing how closely users' experiential evaluations align with the conditions objectively observed in the field. When delivered quality is low but users report high satisfaction, providers may choose not to intervene, although this decision might hinder the ability to attract new users who hold different expectations. Conversely, if delivered quality is high but users remain dissatisfied, further investment in that attribute may be ineffective, as satisfaction in this case may depend on factors unrelated to service quality.

QG4 quantifies the gap between perceived quality (PUS) and expected quality (PUI). This comparison - expressed through percentages rather than mean satisfaction scores - closes the quality loop with a unified output metric. Using percentages helps mitigate the quality paradox (Friman & Fellesson, 2009), whereby negative gaps may originate either from low satisfaction or from excessively high expectations, making interpretation clearer and more consistent across parameters.

Interpreting gap values remains a methodological challenge, as no universal rule applies. As a general guideline, service providers and policy-makers aim to minimise discrepancies across the four quality dimensions. Positive gaps typically indicate areas where the system performs well, whereas negative gaps signal potential weaknesses

requiring attention. To ensure consistent interpretation, all gaps are computed following the direction of the arrows illustrated in *Figure 23*.

4.3.2 Definition of improvement strategies

Once the four quality gaps have been quantified, the subsequent phase concerns their interpretation and prioritisation. This step is essential, as it links the analytical framework to decision-making processes and guides the definition of measures to be undertaken by the various actors involved in the provision and governance of multimodal urban mobility systems: public transport operators, shared micromobility providers, and public authorities.

The prioritisation of gaps does not follow a universal rule; rather, it depends on the strategic objectives of the actors involved. If the aim is to ensure compliance with service certifications, contractual obligations, or benchmarking standards, priority may be given to QG2, which assesses the alignment between designed and delivered quality. Conversely, organisations pursuing a strongly user-centred approach may focus first on QG1 and QG3, as these gaps highlight mismatches between user expectations and both planned and experienced service levels. When the objective is to improve overall user satisfaction and loyalty, QG4 becomes central, since it reveals the difference between what users expect and what they perceive.

After gaps are prioritised, the next step is to define appropriate measures. These interventions can be grouped into preventive and corrective measures.

Preventive measures refer to actions aimed at preserving the current level of service quality in those criteria where the system shows stable or near-acceptable quality. Their purpose is to prevent deterioration over time by maintaining existing infrastructure conditions, ensuring the continuity of functional elements, and avoiding the emergence of new deficiencies. These measures typically include routine maintenance, monitoring, and safeguarding of facilities or services that are already compliant.

Corrective measures, by contrast, are targeted interventions designed to resolve misalignments detected through the Quality Gap Analysis. They address criteria where expected, designed, delivered, and perceived quality diverge substantially. Corrective actions generally involve structural or operational improvements, such as infrastructure upgrades, enhanced service delivery, or redesigning deficient elements, to restore alignment between user needs, planning objectives, and actual system performance.

The selection of measures should account for operational feasibility, financial constraints, and the distribution of responsibilities across stakeholders. As an example, persistent issues in delivered quality (QG2) may require infrastructure upgrades led by public authorities, operational adjustments by transport operators, or contractual revisions involving regulating bodies. In contrast, gaps related to perceived quality (QG3 and QG4) may call for communication initiatives, user-engagement strategies, or targeted

improvements in customer-facing elements such as information clarity, wayfinding, or perceived safety.

A continuous feedback loop is required to monitor the effects of the adopted measures. Measures should be documented, evaluated, and reassessed based on updated quality measurements, ensuring that the system evolves toward progressively better integration, higher quality, and improved user satisfaction.

5. RESULTS AND EMPIRICAL APPLICATIONS

This chapter presents the empirical application of the methodological framework developed in Chapter 4 and illustrates the results obtained across its three main phases: (i) the quality characterisation and the selection of KQPs, (ii) the quality measurement through objective and subjective approaches, and (iii) the quality management based on the analysis of the four quality dimensions and their associated gaps.

The first phase establishes the conceptual foundations for quality assessment in multimodal transport systems. Through a systematic literature review, an expert-based screening process, and the application of the AHP, a set of KQPs is identified, ensuring coverage of the physical, operational, technological, social, and psychological dimensions that shape user experience across different transport modes, including LPT, non-motorised mobility, and electric micromobility.

The second phase operationalises the quality framework through two complementary measurement techniques. Objective quality, corresponding to delivered quality, is assessed using the Mystery Client (MC) technique. This technique was partially tested with a large-scale metropolitan environment (Melbourne, Australia) and subsequently tested comprehensively with a medium-sized European city (Verona, Italy). The MC technique enables the systematic observation of real operating conditions along multimodal paths and origin-destination pairs, translating the selected KQPs into measurable indicators aggregated across quality criteria. Subjective quality, encompassing expected and perceived quality, is measured through a Customer Satisfaction (CS) technique administered to residents of Verona, capturing both the importance attributed to each KQP and the satisfaction associated with users' actual travel experience.

The third phase focuses on quality management and uses the four-dimensional quality loop - expected, designed, delivered, and perceived quality - to identify and quantify quality gaps. Therefore, it was applied only to the Verona case, for which the CS technique was also developed. These gaps are interpreted to highlight strengths, misalignments, and critical issues across modes, criteria, and paths. The results support the distinction between preventive-oriented measures, where service quality meets expectations, and corrective-oriented measures, where gaps reveal inadequacies in planning, provision, or user experience. This phase demonstrates how the integrated framework can guide evidence-based strategies for enhancing the quality and user-centredness of multimodal transport systems.

Overall, the chapter provides a comprehensive empirical validation of the proposed methodology and illustrates its applicability in diverse urban contexts, showing how conceptual, operational, and managerial insights can converge to inform more effective quality governance in sustainable and multimodal mobility systems.

The methodological framework demonstrated in this chapter is intentionally designed to be modular and scalable. Its structure - based on quality characterisation, multilevel measurement, and gap-oriented management - allows the approach to be applied across different urban contexts, transport modes, and spatial scales, from individual corridors to entire multimodal networks. As such, the methodology provides a flexible tool that can support diverse planning, operational, and policy needs in cities of varying size, governance structures, and mobility ecosystems.

5.1 Characterization of quality and Selection of KQPs

This section presents the results of the first phase of the research methodology, which focuses on characterising transport quality and identifying the KQPs that underpin its evaluation. It illustrates the process through which the most relevant KQPs were selected and prioritised, forming the conceptual and operational foundation for the subsequent phases of quality measurement and management.

5.1.1 Identification of KQPs from the literature

According to the first phase of the methodology, the long list of KQPs for a multimodal urban transport system has been derived from the systematic literature review presented in Chapter 2. Moreover, this list was clustered according to eight readjusted criteria from the service quality standards EN 13816:2002 (CEN/TC 320, 2002) and EN 15140:2006 (CEN/TC 320, 2006). *Figure 24* summarizes the papers that emerged from the application of the PRISMA Statement in relation to the quality criteria of European legislation. In addition, the figure illustrates how the identified quality criteria were addressed across different transport macro-categories (e.g., e-micromobility, LPT, non-motorised mobility), specific transport modes or combinations of multiple modes, and across the types of elements analysed (service, infrastructure, or both). This type of mapping also provides a first indication of how frequently each quality criterion has been addressed across specific transport modes and, crucially, in multimodal configurations. By visualising the correspondence between criteria, mode categories, and the scope of analysis (service, infrastructure, or both), the figure helps identify which dimensions of quality have received greater attention in the literature and which remain comparatively underexplored. This is particularly relevant when studies consider more than one mode simultaneously: analysing how criteria are treated in such multimodal contexts allows for a preliminary understanding of how the literature conceptualises the interaction between modes, the extent to which systemic or cross-modal quality aspects are acknowledged.

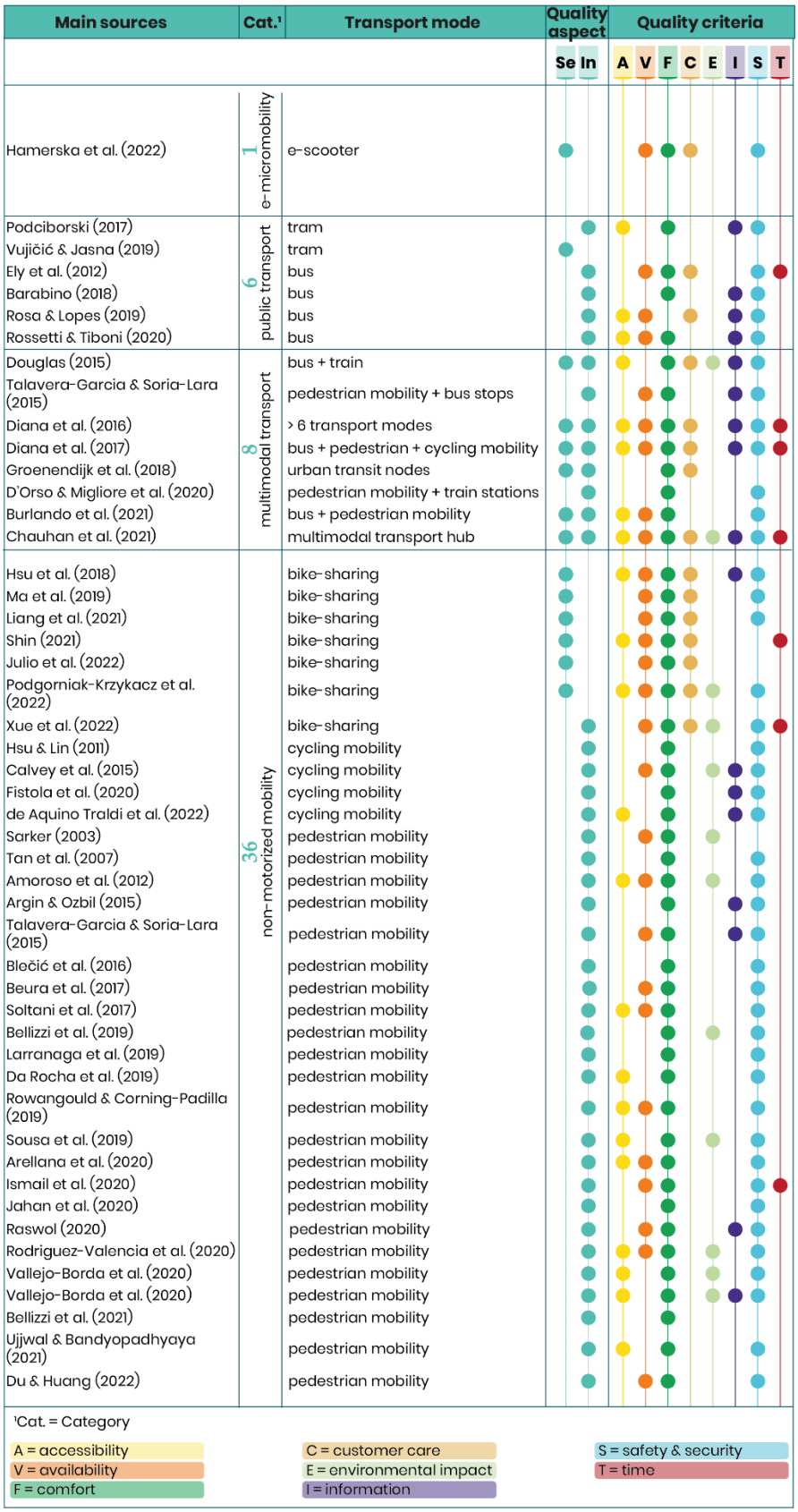


Figure 24. Summary of papers that emerged from the application of the PRISMA Statement in relation to the criteria derived from the service quality standards EN 13816:2002 and and EN 15140:2006. The papers grouped according to the transport mode they address and are presented in chronological order, from the oldest to the most recent.

The extensive list of KQPs is organized into multiple parameter levels. The first level consists of 23 parameters that define individual criteria. These are further refined at the second level, which includes 65 quality parameters (*Table 3*). The second-level parameters are subsequently assessed by academic experts. It is important to note that these parameters have been standardised in terms of terminology, consolidating those expressed differently across different papers but with the same meaning. Therefore, in the papers they come from, the parameters might be called slightly differently and may not be directly related to the eight quality criteria.

Each parameter was then associated first with a quality criterion derived from quality-of-service evaluation methodologies, then with one or more transport systems that it can potentially evaluate (*Table 8*).

Table 8. A long list of criteria and parameters for the assessment of transport quality in the urban environment.

Criteria	Parameter – level I	Parameter – level II	LPT	Pedestrian mobility	Cycling mobility	E-MM
Time	Adherence to schedule / Reliability	Efficiency	●			
		Punctuality	●			
		Regularity	●			
	Length of trip time	At Origin/boarding & alighting/Destination points	●	●	●	●
		Time in-vehicle	●		●	●
		Time for trip planning	●		●	●
Safety & Security	Emergency management	Facilities and plans	●	●	●	●
	Freedom from accidents	Crashes	●	●	●	●
		Avoidance/visibility of hazards	●		●	●
		Travel behaviour	●			
		Presence/visibility of supports, e.g., handrails	●	●	●	●
		Preventive design	●	●	●	●
		Speed	●	●	●	●
		Traffic flow	●	●	●	●
	Freedom from crime	Lighting	●	●	●	●
		People (mis)behaviour	●	●	●	●
		Public safety monitoring	●	●	●	●
Information	General information	Info. about accessibility	●	●	●	●
		Info. about availability	●	●	●	●
		Info. about security	●	●	●	●
		Info. about source of information	●	●	●	●
		Info. about time	●			
		Info. about comfort	●		●	●
		Info. about environmental impact	●		●	●
		Info. about customer care	●		●	●
Environmental impact	Pollution and health	Exhaust	●	●	●	●
		Natural resources	●	●	●	●
Customer care	Assistance	Customer orientation	●		●	●
	Commitment	Innovation and initiative	●		●	●
	Customer interface	Complaints	●		●	●
		Tools and related qualities	●		●	●
	Staff	Staff appearance	●			
		Staff availability	●			
		Staff skills (behaviour)	●		●	●
	Ticketing options	Fare structure and policy	●		●	●
		Consistent price calculations	●		●	●
Comfort	Ambient conditions	Atmosphere	●	●	●	●
		Cleanliness	●	●	●	●
		Disruption	●	●	●	●

		Density		●	●	●
		Visual pollution	●	●	●	●
		Odour	●	●	●	●
		Noise	●	●	●	●
		Weather protection	●	●	●	●
	Complementary facilities	Commercial activities and services	●	●	●	●
		Environment facilities	●	●	●	●
		Entertainment	●			
		Amenities	●	●	●	●
	Ergonomics	Ease of movement			●	●
		Furniture design	●	●	●	●
	Ride comfort	Driving	●			
	Path comfort	Walking and cycling		●	●	●
	Seating and personal space for public transport vehicles	In-vehicle	●			
		At b/a points	●	●		
Availability	Network	Modes	●	●	●	●
		Area covered	●	●	●	●
		Distance to O/D points	●	●	●	●
	Operation	Frequency	●			
		Operating hours	●			
		Load factor	●		●	●
Accessibility	External interface	Interchange/multimodality	●	●	●	●
	Internal interface	Entrance/exits	●	●	●	●
		Internal movement	●	●	●	●
	Ticketing availability	Ticketing acquisition	●		●	●
		Ticketing validation	●			

5.1.2 Expert judgment and evaluation of KQPs

In phase 2, academic expert judgment on the quality of components, attributes, and parameters was elicited through a web-based international survey titled *"Research on Key Quality Parameters of an Integrated and Sustainable Urban Transportation System"*. The targeted panel consisted of academic experts identified through a systematic literature review on quality assessment in urban transport systems. This ensured that the respondents had proven technical and research experience across the full spectrum of the topic, including public transport, non-motorised mobility, and micromobility. Therefore, all authors cited in *Figure 24* were contacted as their academic background and domain knowledge made them suitable to assess the KQPs investigated in the study.

To support the validity of the survey, all questions were pre-tested and piloted before full deployment. Additionally, detailed explanations of the research framework, components, and attributes were provided within the survey to mitigate misunderstandings and reduce potential bias. The clarity of the material, combined with the academic background of participants, further supports the reliability of the collected judgments.

The web survey was structured into two parts. In the first part, they were given two first-level components and asked to indicate which one they considered relevant and to what extent, choosing a value to show the strength of their preference through a pairwise comparison (Appendix C, *Figure C.1*). Then, as with the first question, they were asked to indicate attributes they considered relevant and to what extent (Appendix C, *Figure C.2*).

They compared the two components of Methodological Quality and Quality Relevance and the related ten attributes, according to a 1-9-point Saaty scale, where:

- 1 represents an equal relevance of the compared elements;
- 3 represents a medium relevance, where the judgment slightly favours one element over the other;
- 5 represents a strong relevance, where the judgment strongly favours one element more than the other;
- 7 represents a very strong relevance, where one element is much more strongly favoured than the other;
- 9 represents the overall relevance, where one element is maximally more relevant;
- 2, 4, 6, and 8 represent intermediate values between the previously defined assessments; these intermediate values are useful if judgments do not fit the relevance scale at the previous five levels.

In the second part, experts were asked to rate each parameter on a 1-to-5 scale, assigning a value to each attribute for each parameter of the original list of KQPs (Appendix C, *Figure C.3*).

The survey required three waves of data collection that were carried out between June and September 2024. For the first part, the survey was sent via email to 258 recipients across 34 countries, with 220 successful deliveries (85%) and a total of 27 complete responses collected during the first wave (response rate of 12%). A reminder was sent two weeks after the first wave, and with the second wave, another 18 responses were obtained, for a total of 45 (response rate of 20%).

While this rate may appear modest, it is consistent with typical response rates in similar academic expert surveys, especially in niche fields (Barabino et al., 2020). All respondents possessed recognized expertise in the quality assessment of one of the mobility types considered, ensuring the relevance and quality of their input. Of the respondents, 67% were European (including Italy, the Netherlands, the UK, Spain, Portugal, Belgium, Greece, and Croatia), 17% were from North and South America, and 11% were from Asia. While responses were concentrated mainly in Europe, this is influenced by the panel of studies selected, which guided the identification of experts. Specifically, the comparison between the panel and respondent distributions shows coherence for North and South America, a slight overrepresentation for Europe, and an underrepresentation for Asia.

Table 9 provides an overview of the distribution of email invitations, deliveries, and responses by continent.

Table 9. Composition of the Sample and Survey Participation Rates.

Country	Phase 1: waves 1 and 2			
	Email sent	**Email delivered**	**Responses received**	**Response rate (%)**
Europe	145	122	33	27
North & South America	47	45	7	16
Asia	62	50	5	10
Oceania	4	3	0	0
Total	**258**	**220**	**45**	**20**

To ensure a timely progression of the research and to deepen the analysis, the second part of the survey was distributed in October 2024 to a selected group of academic experts with specialized knowledge of the topic. This targeted approach facilitated more direct engagement with high-quality and consistent feedback that contributed to the robustness of the findings.

Next, the judgments of academic experts were analysed. They were disaggregated and systematized using a matrix system in Excel, where each academic expert (identified by an ID number) was associated with one matrix (Appendix B, *Figure B.1* and *Figure B.2*).

Thus, the AHP has been applied to calculate the weights of components and attributes, and their average according to the Eqns 1-5, which are described in the methodology section. In analysing evaluations of the pairwise comparisons between the attributes, only four evaluations with consistency values strongly above 1 (thus, inconsistent) were deleted. Indeed, these were probably ratings made without substantial criteria. Therefore, all 45 judgments were retained for the pairwise comparison of the components, while the original ratings decreased from 45 to 41 for the pairwise comparison of the attributes.

Then, Eqns 6-12 were applied to obtain the WKQPS. The average weight (μ), standard deviation (σ), and coefficient of variation (σ/μ) of components and attributes are presented in *Table 10*. While the standard deviation (σ) values are almost the same for the Methodological quality and Quality relevance, the aggregate average weight (μ) values differ slightly, showing a preference for the Methodological quality. However, the coefficient of variation (σ/μ) is higher for the Quality relevance and shows a slight spread among the responses, likely due to greater subjective interpretation and contextual influence, which might depend on local priorities, project objectives, and disciplinary perspectives.

Table 10. Average weights of the components and attributes.

Components	Aggregate average weight (μ)	Standard deviation (σ)	Coefficient of variation % (σ/μ)
Methodological quality	0.513	0.260	50.8
Quality relevance	0.487	0.267	54.8
Methodological quality attributes			
Measurability	0.262	0.099	37.8
Easy Availability	0.200	0.128	64.1
Speed Availability	0.090	0.033	36.2
Interpretability	0.231	0.093	40.1
Relevance to multiple transport systems	0.218	0.130	59.6
Quality relevance attributes			
Integration between stakeholder and user perspectives	0.147	0.098	66.5
User-orientation	0.193	0.112	58.3
Measurability from subjective and objective perspectives	0.170	0.091	53.6
Quantification of the number of users	0.208	0.124	59.4
Relation between transport and public space qualities	0.282	0.117	41.5

Academic experts considered Interpretability ($\mu \simeq 0.231$) and Measurability ($\mu \simeq 0.262$) as the most relevant attributes of Methodological quality, suggesting a preference for parameters that allow robust, objective assessment. In contrast, the Easy Availability and Speed Availability showed variability that highlights context-specific influences, i.e., contexts with accessible data vs. those with informational constraints. Within the attributes of Quality relevance, the most relevant ones were the Relation between transport and public space qualities ($\mu \simeq 0.282$) and Quantification of the number of users ($\mu \simeq 0.208$). These results highlight the high value of functional urban integration and tangible usage metrics. In contrast, the lower weight of the Integration between stakeholder and user perspectives ($\mu \simeq 0.147$) and high variability (66.5%) may reflect the differing institutional practices or difficulties coordinating multi-actor perspectives. Medium values characterise the User-orientation ($\mu \simeq 0.193$) and Measurability from subjective and objective perspectives ($\mu \simeq 0.170$).

5.1.3 Assessment and prioritization of KQPs

In phase 3, a first WKQPS is obtained by additive weighting, aggregating the outcome weights and scores to evaluate the performance of each KQP, and providing a preliminary selection of the most promising KQPs. However, since the sample of responses to the web survey was small, a probability of bias was admitted. Therefore, the Monte Carlo simulation was applied to mitigate it. This technique permits the production of artificial data that needs to be matched with the observed data based on academic expert judgments, deriving the Adjusted Weighted Key Quality Parameter Score (AWKQPS) for each KQP. Matching is achieved by using a probability distribution that most accurately reflects the current state of knowledge. In this study, the Monte Carlo simulation is applied to the set of KQPs related to each attribute. The observed distribution of the marks given by academic experts is simulated 1,000 times for each attribute (thus 10,000 times for

each parameter) in the [1,5] interval. The score of each parameter (based on 650,000+ Monte Carlo simulations) was computed by applying Eqn. 12, which returned the final scores for each parameter. *Table 11* shows the two scores, i.e., WKPQS and AWKPQS as well their variation for each parameter, following the descending order of the AWKPQS.

Table 11. Final ranking of Weighted Key Parameters Quality Scores.

n	Criteria	Parameter I	Parameter II	WPQS	AWKPQS	Score variation
1	Comfort	Seating and personal space for public transport vehicles	At b/a points*	4.297	4.302	=
2	Availability	Operation	Frequency*	4.197	4.227	-
3	Safety & Security	Freedom from accidents	Traffic flow	4.233	4.221	+
4	Accessibility	External interface	Interchange / multimodality	4.128	4.157	+
5	Comfort	Seating and personal space for public transport vehicles	In vehicle*	4.150	4.147	+
6	Availability	Operation	Operating hours*	4.031	4.075	-
7	Availability	Network	Modes	4.143	4.069	-
8	Availability	Network	Area covered	4.058	4.038	-
9	Comfort	Ambient conditions	Density	3.909	3.981	-
10	Time	Adherence to schedule / Reliability	Punctuality*	4.022	3.934	+
11	Time	Length of trip time	At Origin/boarding & alighting/Destination points	3.913	3.915	-
12	Safety & Security	Freedom from accidents	Crashes	3.908	3.908	=
13	Accessibility	Ticketing availability	Ticketing acquisition	3.847	3.853	=
14	Comfort	Ambient conditions	Cleanliness	3.832	3.834	=
15	Availability	Network	Distance to O/D points	3.746	3.793	+
16	Time	Length of trip time	Time in-vehicle	3.796	3.781	=
17	Safety & Security	Freedom from crime	Lighting	3.809	3.700	+
18	Safety & Security	Freedom from accidents	Speed	3.708	3.768	+
19	Accessibility	Internal interface	Entrance/exits	3.748	3.750	-
20	Accessibility	Ticketing availability	Ticketing validation*	3.769	3.732	+
21	Availability	Operation	Load factor	3.745	3.718	-
22	Information	General information	Info. about time*	3.746	3.715	=
23	Comfort	Complementary facilities	Amenities	3.777	3.715	-
24	Comfort	Complementary facilities	Commercial activities and services	3.775	3.715	+
25	Comfort	Ergonomics	Ease of movement*	3.668	3.699	-
26	Comfort	Path comfort	Walking and cycling	3.702	3.697	=
27	Information	General information	Info. about accessibility**	3.709	3.684	-
28	Customer care	Customer interface	Complaints**	3.587	3.655	-
29	Information	General information	Info. about availability	3.598	3.635	-
30	Comfort	Complementary facilities	Environment facilities	3.658	3.624	+
31	Comfort	Ambient conditions	Disruption	3.617	3.623	-
32	Customer care	Ticketing options	Fare structure and policy	3.650	3.621	+
33	Safety & Security	Freedom from accidents	Presence/visibility of supports, e.g. handrails	3.574	3.606	=
34	Time	Adherence to schedule / Reliability	Regularity*	3.529	3.572	=
35	Information	General information	Info. about security	3.515	3.488	=
36	Comfort	Ambient conditions	Weather protection	3.494	3.474	=
37	Comfort	Ambient conditions	Noise	3.449	3.449	=

n	Criteria	Parameter I	Parameter II	WPQS	AWKPQS	Score variation
38	Information	General information	Info. about comfort	3.441	3.421	=
39	Comfort	Complementary facilities	Entertainment*	3.389	3.402	-
40	Customer care	Ticketing options	Consistent price calculations	3.408	3.389	+
41	Accessibility	Internal interface	Internal movement	3.359	3.326	=
42	Comfort	Ride comfort	Driving*	3.346	3.296	=
43	Safety & Security	Freedom from crime	Public safety monitoring	3.279	3.295	-
44	Information	General information	Info. about customer care	3.223	3.280	-
45	Comfort	Ergonomics	Furniture design	3.308	3.265	+
46	Time	Length of trip time	Time for trip planning	3.300	3.259	+
47	Safety & Security	Freedom from crime	People (mis)behaviour	3.221	3.250	=
48	Information	General information	Info. about source of information	3.216	3.238	=
49	Customer care	Assistance	Customer orientation	3.154	3.183	=
50	Time	Adherence to schedule / Reliability	Efficiency*	3.112	3.103	=
51	Comfort	Ambient conditions	Odour	3.043	3.058	+
52	Safety & Security	Emergency management	Facilities and plans	2.991	3.020	-
53	Safety & Security	Freedom from accidents	Preventive design	2.964	3.018	-
54	Safety & Security	Freedom from accidents	Avoidance/visibility of hazards	2.974	3.001	-
55	Information	General information	Info. about environmental impact	3.058	2.997	+
56	Comfort	Ambient conditions	Visual pollution	3.002	2.992	+
57	Safety & Security	Freedom from accidents	Travel behaviour*	3.022	2.955	+
58	Customer care	Customer interface	Tools and related qualities	2.982	2.899	-
59	Customer care	Staff	Staff availability*	2.883	2.888	=
60	Customer care	Commitment	Innovation and initiative	2.880	2.852	=
61	Customer care	Staff	Staff skills (behaviour)	2.817	2.822	-
62	Comfort	Ambient conditions	Atmosphere	2.820	2.799	-
63	Environmental impact	Pollution and health	Natural resources**	2.743	2.796	-
64	Environmental impact	Pollution and health	Exhaust	2.763	2.709	+
65	Customer care	Staff	Staff appearance*	2.582	2.644	=

* Those highlighted in yellow are the parameters that do not evaluate at least 3/4 modes of transport included in this research, which, therefore, are excluded from the ranking.

** Those highlighted in green are the parameters that belong to the criteria excluded from the first quartile, which, therefore, were added to the final ranking.

As can be seen from the selection of the total AWKPQS, in the first quartile, there are no parameters that belong to the criteria of Information, Customer care, and Environmental impact. Then, a parameter for each of these three criteria was added, choosing those with the highest value in the total score. Specifically, the parameters of Information about accessibility, Complaints, and Natural resources were added. At this point, the final set of KQPs consists of 19 quality parameters. However, not all of them can evaluate at least three out of four transport systems included in the research (*Table 8*). Specifically, five parameters in the first quartile do not meet this selection criterion and were replaced with five other parameters with the highest value in the total score. The parameters replaced

are: At b/a points, Frequency, In-vehicle, Operating hours, and Punctuality. *Table 12* shows the final selection of KQPs.

Table 12. Final selection of KQPs following the order of the AWKPQS.

n	Criteria	Parameter – level I	Parameter – level II	LPT	Pedestrian mobility	Cycling mobility	E-MM
1	Safety & Security	Freedom from accidents	Traffic flow	●	●	●	●
2	Accessibility	External interface	Interchange / multimodality	●	●	●	●
3	Availability	Network	Modes	●	●	●	●
4	Availability	Network	Area covered	●	●	●	●
5	Comfort	Ambient conditions	Density	○	●	●	●
6	Time	Length of trip time	At Origin/boarding & alighting/Destination points	●	●	●	●
7	Safety & Security	Freedom from accidents	Crashes	●	●	●	●
8	Accessibility	Ticketing availability	Ticketing acquisition	●	○	●	●
9	Comfort	Ambient conditions	Cleanliness	●	●	●	●
10	Availability	Network	Distance to O/D points	●	●	●	●
11	Time	Length of trip time	Time in-vehicle	●	○	●	●
12	Safety & Security	Freedom from crime	Lighting	●	●	●	●
13	Safety & Security	Freedom from accidents	Speed	●	●	●	●
14	Accessibility	Internal interface	Entrance/exits	●	●	●	●
15	Availability	Operation	Load factor	●	○	●	●
16	Comfort	Complementary facilities	Amenities	●	●	●	●
17	Information	General information	Info. about accessibility	●	●	●	●
18	Customer care	Customer interface	Complaints	●	○	●	●
19	Environmental impact	Pollution and health	Natural resources	●	●	●	●

5.1.4 Discussion and methodological implications

Table 12 shows that among the parameters belonging to the first quartile of the complete list, there is practically no difference between WKPQS and the AWKPQS, which share 15 of 16 KQPs. The Monte Carlo simulation enhanced the scores of each parameter without altering their relative importance, though the adjustment of the original scores led to minor changes in the ranking of some parameters. Furthermore, the use of a set of random numbers as input to iteratively assess the deterministic model turned the observed model into a stochastic one. Dwelling on the first quartile, the most important quality criteria are Safety and Security and Availability, followed by Comfort, Accessibility, and Time. In contrast, Customer Service, Information, and Environmental Impact are considered less important. The most relevant criteria of Safety and Security, Availability, Comfort, Accessibility, and Time are applicable across all transport systems, including walking, cycling, public transport, and shared mobility. Their universality likely explains their higher scores, as they reflect core expectations shared by all users. In contrast, criteria such as

Customer Service and Information are tied to specific modes and interaction formats, and are not relevant to all systems, particularly pedestrian mobility. Environmental impact, instead, although crucial from a planning perspective due to the sector's contribution to externalities, appears less directly linked to perceived quality, highlighting a disconnect between user perceptions and the environmental implications of transport systems.

To further investigate the relevance of the best KQPs, we compared the results with ranking methods, evaluating whether the 19 best-ranked KQPs are considered equally important in the surveyed literature conducted with the PRISMA statement (Chapter 2). The former ranking method consists of a Bibliometric Score, i.e., how many times the paper that used each indicator has been cited. The latter ranking method consists of a Citation Score, i.e., a count of how many times each parameter appears in the literature analysed (i.e., maximum once per paper).

Although a careful comparison could not be performed since some parameters in the literature refer to a single mode of transport, and a similar study has not yet been performed, Bibliometric Score and CS show similar results to each other but are quite different from those returned by AWKQPS based on academic expert judgments and Montecarlo's adjustment. Compared with the AWKQPS, the strong preponderance of comfort parameters in the Bibliometric Score and Citation Score s is particularly noticeable. In addition, both Bibliometric Score and Citation Score include parameters belonging to the criteria of Customer care and Environmental impact, which are missing in the AWKPQS. The Bibliometric Score shares only one Safety and Security parameter with the AWKPQS (Speed), while the Citation Score shares three (Interchange/multimodality, Modes, At O/boarding and alighting/D points) corresponding to the criteria of Accessibility, Availability, and Time.

Figure 25 shows the seven parameters that overlap in all three scores, which collectively form the final set of KQPs: Information about accessibility, Amenities, Cleanliness, Area covered, Entrance/exits, Lighting, and Traffic flow. The final ranking includes two parameters for Safety and Security, two for Comfort, one for Availability, one for Accessibility, and one for Information. Environmental Impact, Customer Care, and Time are excluded from the ranking. The absence of parameters for Environmental Impact and Customer Care is coherent with the initial ranking, while a parameter for Information (Info. about accessibility), which was initially missing, appears, and Time is removed from the ranking.

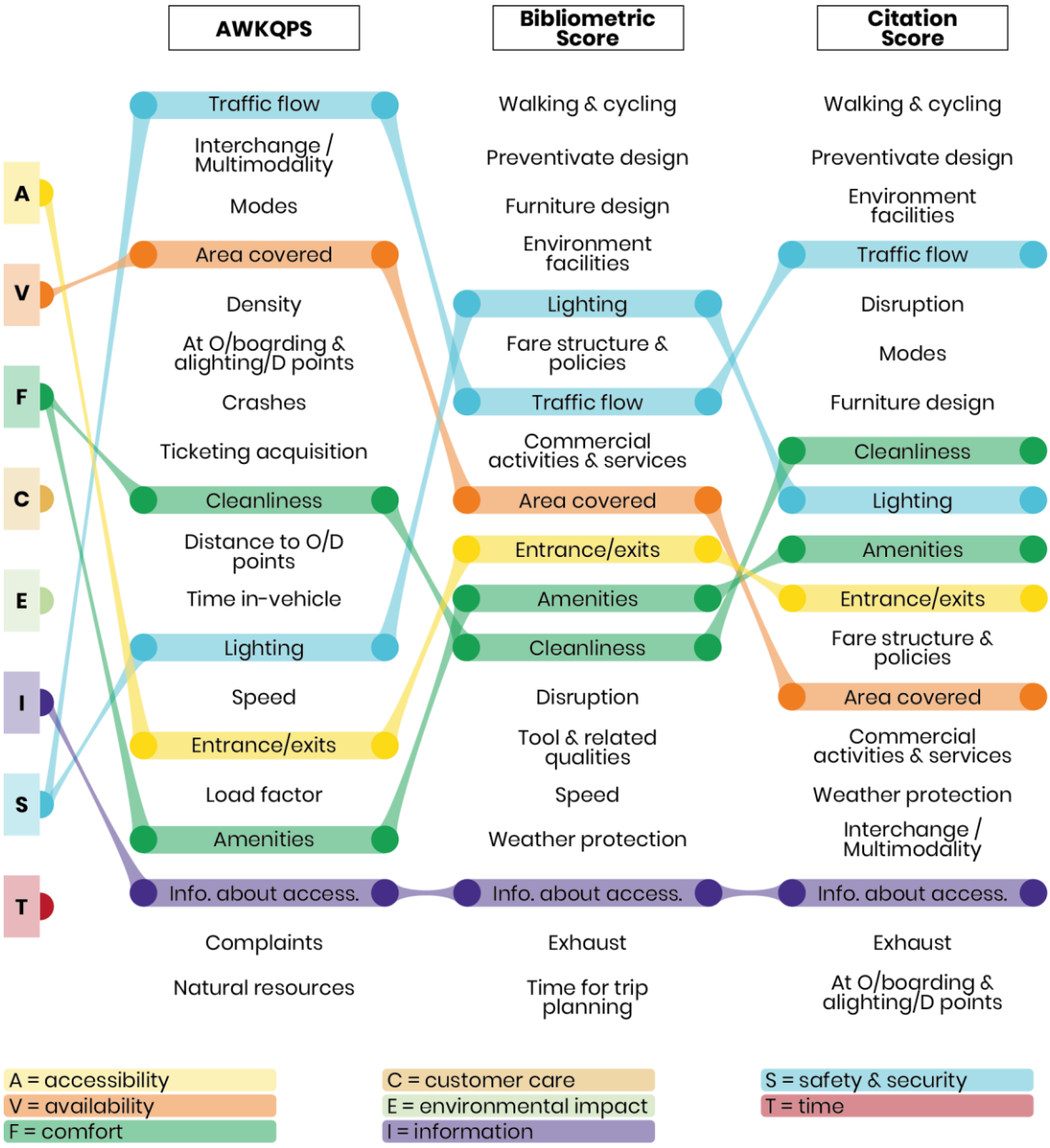


Figure 25. Final set of KQPs ranked by AWKQPS, Bibliometric and Citation Scores. The final set of the best KQPs is given by the seven parameters that are in common among the three scores used.

This final set of quality parameters emphasises the intrinsic link between urban design and transport planning, highlighting the role of urban elements in enabling or constraining mobility. Literature consistently shows that urbanistic features such as well-designed sidewalks, barrier-free bus stops, and ramps play a crucial role in shaping accessibility and comfort (Rosa & Da Cruz Lopes, 2019; Podciborski, 2017; Douglas, 2015). The internal accessibility of transport modes, reflected in parameters like entrances and exits, is not only a matter of compliance with universal design principles (Chauhan et al., 2021) but also a determinant of seamless multimodal integration, as evidenced by the importance of dedicated access points for public transport and bike-sharing systems (Burlando et al., 2021; Shin, 2020). Availability further extends to spatial continuity and coverage, where

the regularity of pedestrian and cycling infrastructure critically affects user experience and mode choice (Rossetti, 2020; Liang et al., 2021). Comfort is closely tied to the maintenance and cleanliness of these urban infrastructures, a factor linked to overall satisfaction (Barabino, 2018; Arellana et al., 2019; Sousa et al., 2019). Amenities such as benches, public restrooms, and drinking fountains contribute to comfort by addressing basic human needs during travel, thus enhancing the attractiveness of active and shared mobility options (Soltani et al., 2017; Groenendijk et al., 2018; Ujjwal & Bandyopadhyaya, 2021). Information, particularly related to accessibility, emerges as a key enabler for all users, but especially for those with mobility restrictions, emphasising the role of clear signage and inclusive information design in facilitating independent travel (de Aquino Traldi et al., 2022; Diana et al., 2016). Safety and security are critical factors influencing both the actual risk of accidents or crime and users' perception of travel quality (Tan et al., 2007; Fistola, 2020). These dimensions affect mobility choices and comfort, underscoring the importance of designing urban environments and transport services that minimise hazards. Closely related to these aspects is the management of traffic flow, which involves balancing the interactions between pedestrians, cyclists, and motor vehicles to reduce conflicts and enhance overall safety (Rodriguez-Valencia et al., 2020; Ma et al., 2019; D'Orso & Migliore, 2019). Finally, lighting plays a fundamental role in safety, particularly in public transport stops and pedestrian paths during evening or night-time hours, contributing to a sense of security and usability in urban spaces (Jahan et al., 2020; Soltani et al., 2017; Rowangould & Corning-Padilla, 2019).

In addition, the final set of KQPs covers all thematic clusters identified in the literature, confirming the representativeness of the selection process. The parameter Entrances/Exits reflects the Geometric & Physical cluster, as it captures spatial configuration and design features essential for accessibility. Area Covered corresponds to the Operational cluster, highlighting the importance of network continuity and service availability. Amenities and Cleanliness relate respectively to the Geometric & Physical and Maintenance & Conservation clusters, integrating both infrastructural provision and upkeep, representing such a broad criterion as comfort. The Customer Service parameter aligns with the Operational cluster, emphasising service delivery and customer care. Natural Resources falls within the Environmental & Aesthetic cluster, capturing sustainability aspects associated with environmental impact. Traffic Flow and Lighting correspond respectively to the Operational and Social & Psychological clusters, in reference to safety and security within the urban environment. Finally, the length of the trip is associated with the Operational cluster, reflecting system performance in terms of efficiency. Together, these correspondences demonstrate that the selected KQPs capture the full spectrum of dimensions identified in the literature, ensuring a balanced and multidimensional representation of multimodal transport quality.

Overall, these parameters illustrate how quality in multimodal urban mobility cannot be fully understood without considering the spatial and experiential aspects of the urban environment. The interplay between infrastructure design, service provision, and user

needs requires holistic planning that integrates transport with urban space to foster high-quality, sustainable mobility.

The final set of KQPs offers a practical framework for assessing both the objective and perceived quality of multimodal urban transport systems. These parameters can support public authorities and transport providers in identifying critical issues across the full travel chain, from origin to destination, encompassing infrastructure, service provision, and user experience. Unlike traditional evaluations focused on individual modes or segments, this approach enables a holistic understanding of mobility quality, guiding targeted improvements and more integrated, user-centred planning.

5.2 Quality measurement in multimodal transport systems

The second phase of the research operationalises the methodological framework developed in Chapter 4 by applying it to real-world multimodal transport environments. This phase focuses on the empirical measurement of quality through the integration of objective and subjective data, enabling a comprehensive evaluation of how mobility systems perform under actual operating conditions and how this quality is perceived by users. Delivered quality is assessed using the Mystery Client (MC) technique, which captures the quality of services and infrastructures as experienced in situ across different modes and urban contexts. Expected and perceived quality are assessed through Customer Satisfaction (CS) technique, which provide insight into the importance attributed by users to KQPs and their satisfaction with the corresponding service attributes.

The combined use of these complementary data sources allows for a multidimensional and user-centred assessment of system quality, supporting the identification of critical gaps between designed, delivered, expected, and perceived quality. This integrative approach is then applied to two distinct empirical settings: Melbourne (Australia) and Verona (Italy), which differ in terms of system complexity, operational conditions, and methodological scope. Melbourne serves as an initial application of the objective component of the framework, focusing on delivered quality through MC observations, while Verona enables the full implementation of the methodology by encompassing both MC and CS assessments. The results of these case studies provide empirical evidence of the framework's applicability across diverse contexts and scales, illustrating its capacity to support robust, comparative, and actionable quality analysis in multimodal urban mobility systems.

5.2.1 Case study I: Melbourne, Australia

The city of Melbourne, Australia, constitutes a Local Government Area located at the core of the Greater Metropolitan Melbourne Region. Its urban fabric is characterised by a dense distribution of transport interchanges, including several major nodes - primarily associated with metropolitan rail stations - and numerous minor interchanges formed by bus and tram

stops. The northern sector of the Central Business District (CBD) hosts a concentration of high-intensity activity centres, such as universities, hospitals, schools, cultural venues, and recreational spaces. This area is also highly residential (Australian Bureau of Statistics, 2016), generating substantial daily travel flows and making it a particularly suitable context for analysing multimodal mobility interactions.

Given these characteristics, the Melbourne case study focuses on a set of representative interchanges within this area, selected to test the applicability of the proposed methodology for evaluating delivered quality in a multimodal environment. Three metropolitan railway stations -Flagstaff, Melbourne Central, and Parliament - were identified as major interchanges, while three tram stops were selected as minor interchanges: “University of Melbourne/Royal Parade”, “Melbourne University/Swanston Street”, and “Melbourne Museum/Nicholson” (*Figure 26*). These stops were chosen due to their proximity to key generators of mobility demand. Specifically, the first is adjacent to the Royal Melbourne Hospital and the University of Melbourne; the second is located near the Royal Dental Hospital of Melbourne and university facilities; and the third serves St Vincent’s Hospital, Carlton Gardens, the Royal Exhibition Building, and the Academy of Mary Immaculate High School.

By focusing on these interchanges, the case study captures a variety of multimodal interactions and provides an appropriate setting for testing a preliminary, simplified application of the Mystery Client technique introduced in Chapter 4.2.3. This exploratory implementation aims to validate the indicator set, observation protocol, and weighting structure before applying the full conformity-based methodology in a subsequent case study.

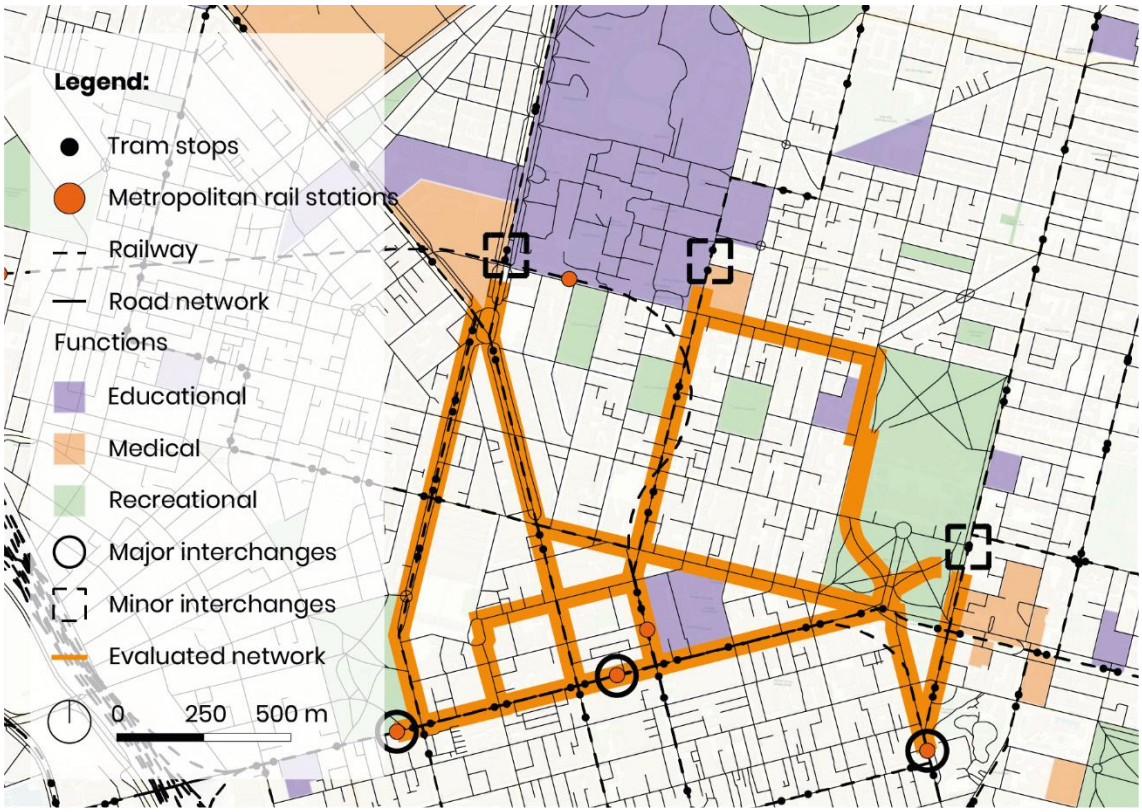


Figure 26. Route framing connecting major with minor interchanges in the area north of Melbourne CBD.

The selected routes do not reflect direct user travel demand but instead represent strategically important segments connecting major and minor interchanges within the study area. This constitutes a key difference from the subsequent case study, which examines the entire travel experience from origin (O) to destination (D). In Melbourne, the focus is placed on assessing the quality of infrastructure and services along these connecting segments, which may form part of broader multimodal trips involving multiple transport modes. By analysing the links between interchanges, rather than the interchanges themselves, the assessment provides insight into the underlying multimodal quality of the system independently of specific travel patterns.

Route selection followed the hierarchy of sustainable mobility (Banister, 2008; Gössling, 2016), which prioritises walking, followed by cycling, public transport, motorcycles, and private vehicles. Consequently, the connecting routes were identified using the shortest pedestrian paths generated through Google Maps. These paths were then adopted as a common reference alignment for other modes (e.g., cycling, public transport, or e-scooters), while acknowledging that actual routes may vary across modes. For example, the shortest cycling path may differ from the pedestrian route, and public transport services may follow alternative street alignments. However, because the goal of this study is not to evaluate mode-specific routing but to assess multimodal quality along predefined segments, these differences are incorporated through appropriate indicators (e.g., presence of infrastructure, service availability, walking time to the nearest stop).

The nine connecting routes linking the three railway stations to the three tram stops were mapped using GIS, and their proximity to major activity nodes reinforced their relevance (*Figure 27).* For each route, the KQPs identified during the second phase of the methodology were measured. Each KQP, associated with a specific quality criterion, was operationalised through one or more indicators drawn from the literature. Indicator selection followed three principles: ease of measurement, rapid data collection, and availability of information through fieldwork or official sources. To ensure alignment with the full set of EN-based criteria, additional KQPs belonging to Environmental Impact, Customer Care, and Time were included, even though they had not been part of the earlier comparison between AWQPS, Bibliometric Score, and Citation Score.

The interchanges analysed here serve as representative examples for testing the measurement framework in its simplified, preliminary form. Nevertheless, the approach is fully scalable and can be extended to a broader set of interchanges, a denser network of connecting routes, or an entire urban mobility system.



Figure 27. Framing of the nine routes analyzed in the Melbourne CBD.

In this part of the research, the data used to evaluate the quality of the connections between interchanges were obtained through both desk-based and field-based analyses. Only 22% of the required information was available through official open-data sources; consequently, 78% of the data was collected through direct field inspections. Fieldwork supported the acquisition of the following information:

- Accessibility indicators, including the presence of ramps at LPT stops and along pedestrian and cycling infrastructure.
- Availability indicators, such as the presence and continuity of LPT, pedestrian, and cycling infrastructure. These represent 71% of the availability-related metrics. The

presence of shared scooters and bikes was assessed using the Lime application, the main shared-mobility provider in the city.

- Comfort indicators, all collected on-site except for the availability of public toilets - obtained from the Department of Health and Aged Care - and the number of drinking facilities, derived from ESRI data.
- Customer care indicators, partly collected in the field (e.g. presence of customer service), while information on hotline availability was sourced from the Victorian Government.
- Safety and Security indicators, such as operational lighting systems, were collected in the field. Traffic-flow data, also contributing to this criterion as well as to environmental impact indicators, were computed in GIS using spatial information from the Victorian Government Data Directory.
- Time-related indicators, including waiting times for LPT services and search times for shared-mobility vehicles. Travel times for different modes were obtained from Google Maps.

Field measurements were carried out on weekdays during February and March 2025, primarily during peak hours (7:00-10:00 a.m. and 5:00-8:00 p.m.). Additional inspections were undertaken after 7:00 p.m. to assess lighting conditions under the Safety and Security criterion. The evaluation of traffic flow required peak-hour measurements, while nighttime observations were necessary to assess perceived safety conditions. Data on non-motorised vehicle flows were retrieved from the Department of Transport and Planning Open Data Hub (DTP, 2020), which provides annual average daily traffic volumes for freeways (excluding toll roads) and arterial roads in Victoria, based on 2020 traffic surveys and estimates.

According to the Eqn. 18, the quality measurements were applied to each of the nine analysed routes, which were subsequently compared to assess their relative quality. The assignment of indicators to the selected KQPs produced a structured set of 43 quality indicators, each associated with a predefined maximum score (*Table 13*).

Table 13. Quality criteria, parameters and indicators, classified based on the score of criteria (min-max).

Criteria	**Parameter – level I**	**Parameter – level II**	**Indicators**	**Max Score**
Customer Care	Assistance	Customer orientation	2	55
	Commitment	Innovation & initiatives		
Environmental impact	Pollution & health	Natural resources	2	60
Time	Length of the trip	At O/b and a/D points	5	85
Accessibility	Internal interface	Entrance/exits	3	105
Safety & Security	Freedom from accidents	Traffic flow	4	125
	Freedom from crime	Lighting		
Comfort	Ambient conditions	Cleanliness	15	160
	Complementary facilities	Amenities		
Information	General information	Info. about accessibility	5	165
Availability	Network	Area covered	7	245
Total			**43**	**1,000**

From *Table 14* to *Table 21* the full set of selected indicators together with their scoring rules, maximum scores, and weighting structure are presented for each criterion.

Accessibility

The criterion of Accessibility generally refers to the access to the transport system (internal interface), including interface with other transport modes (external interface), and the ticketing availability issues, including ticket acquisition and validation (CEN/TC 320, 2002). In this work, we focus on the internal interface first-level parameter related to the second-level parameter of entrance/exits. Specifically, we consider as indicators the presence of ramps at the LPT stops, along the pedestrian and cycling infrastructure (*Table 14*). Ramps represent an element whose presence or absence can influence users' perception of an infrastructure (Ujjwal & Bandyopadhyaya, 2023), as they allow easy access to public transport stops, or to the sidewalk for elderly persons, users with reduced mobility or those travelling with strollers/trolleys; also in the case of cycling, ramps facilitates the access to cycle infrastructure when the street level changes.

Table 14. Indicators for accessibility criterion.

n	Indicators	Scale	Score	Weight	Max Score
1	ramps at LPT stops	not present	0	3.5	35
		present	10		
2	ramps to get on and off to the pedestrian infrastructure	not present	0	3.5	35
		present	10		
3	ramps to enter and exit to the cycling infrastructure	not present	0	3.5	35
		present	10		

Availability

Availability refers to the extent of the transport system offered in terms of geography, time, frequency, and transport mode (CEN/TC 320, 2002). In this work, we focus on the area covered by multiple transport systems in a given network. Specifically, we consider indicators for service/infrastructure presence and continuity of different transport systems (*Table 15*). These types of indicators refer to the presence of LPT, pedestrian and cycling infrastructure, bike-sharing and scooter services, as well as to the degree of continuity of pedestrian and cycling infrastructure. Continuity represents a geometric factor of the network (Rodriguez-Valencia et al., 2020) that can influence the perception of pedestrians and cyclists, including concerning bike-sharing service (Xue et al., 2022).

Table 15. Indicators for availability criterion.

n	Indicators	Scale	Score	Weight	Max Score
4	LPT service/infrastructure	not present	0	3.5	35
		present	10		
5	pedestrian infrastructure	not present	0	3.5	35
		present	5		
		present on both side	10		
6	cycling infrastructure	not present	0	3.5	35
		present	5		
		present on both side	10		
7	bike-sharing service	not present	0	3.5	35
		present	10		
8	scooters-sharing service	not present	0	3.5	35
		present	10		
9	continuity of the pedestrian infrastructure	interrupted	0	3.5	35
		continuous	10		
10	continuity of the cycle infrastructure	interrupted	0	3.5	35
		continuous	10		

Comfort

The criterion of Comfort refers to the elements that make journeys relaxing and pleasurable (CEN/TC 320, 2002). Included in this category are the first-level parameters of Ambient conditions, Complementary facilities, Ergonomics, Ride comfort, Path comfort, Seating and personal space for public transport users. Specifically, we consider two of these parameters: Ambient conditions, represented by the second-level parameter of Cleanliness and the indicators of the presence of bins, trash/broken glasses, and illegal graffiti around; and Complementary facilities, represented by the second-level parameter of Amenities and the indicators of the presence of seating facility, public toilet, bicycle racks, and drinking water facility (*Table 16*). These types of indicators refer to elements related to the cleanliness and good maintenance of infrastructure and urban spaces (i.e. graffiti on walls and buildings), as well related to the provision of complementary facilities that can help during the trip, improving the pleasantness (D'Orso & Migliore, 2019).

Table 16. Indicators for comfort criterion.

n	Indicators	Scale	Score	Weight	Max Score
11	bins at the LPT stops	not present	0	1	10
		present	10		
12	bins along the pedestrian infrastructure	not present	0	1	10
		present	10		
13	trash/broken glasses at the LPT stops	not present	0	1	10
		present	10		
14	trash/broken glasses along the pedestrian infrastructure	not present	0	1	10
		present	10		
15	trash/broken glasses along the cycling infrastructure	not present	0	1	10
		present	10		
16	illegal graffiti at the LPT stops	not present	0	1	10
		present	10		
17	illegal graffiti along the pedestrian infrastructure	not present	0	1	10
		present	10		
18	illegal graffiti along the along the cycling infrastructure	not present	0	1	10
		present	10		
19	seating facility at the LPT stops	not present	0	2	20
		present	10		
20	seating facility along the pedestrian infrastructure	not present	0	0.5	5
		present	10		
21	public toilet	not present	0	1	10
		present	5		
		accessible for people with disabilities	10		
22	bicycle racks at interchanges	not present	0	2	20
		presence of uncovered racks	5		
		presence of covered racks	10		
23	bicycle racks along the route	not present	0	1	10
		present	10		
24	drinking water facility	not present	0	1	10
		present	10		
25	number of drinking water facility	low [at least one every 1000 m]	0	0.5	5
		medium [at least one every 500 m]	5		
		high [at least one every 250 m]	10		

Customer care

The criterion of Customer care refers to the service elements introduced to affect the closest practicable match between the standard service and the requirements of any individual customer (CEN/TC 320, 2002). This category includes the first-level parameters of Assistance, Commitment, Customer interface, Staff, and Ticketing options. Also in this case, we consider two measures: Assistance, represented by the second-level parameter of Customer orientation and by the indicators of the presence of customer service; and Commitment, represented by the second-level parameter of Innovation and initiative and by the indicator of access to a hotline (*Table 17*). These types of indicators refer to the presence and type (physical or virtual) of the customer service working at the public transport stops or for shared-systems services, as well the access to a hotline near public transport stops or along the pedestrian or cycle infrastructure. These aspects are considered important for users not only in relation to LPT, but also to e-micromobility (Hamerska et al., 2022) and bike-sharing systems (Shin, 2020).

Table 17. Indicators for customer care criterion.

n	Indicators	Scale	Score	Weight	Max Score
26	customer service	not present	0	3.5	35
		virtual	5		
		physical	10		
27	access to a hotline	not present	0	2	20
		present	10		

Environmental impact

The criterion of Environmental impact refers to the effect on the environment resulting from a transport system (CEN/TC 320, 2002) and concerns the main first-level parameter of pollution and health. In this work, we consider as a second-level parameter the Natural resources with the indicator of land consumption (*Table 18*). Specifically, we focus on the presence or absence of natural surfaces, such as green areas and urban parks, and their dimension expressed in square meters (m^2).

Table 18. Indicators for environmental impact criterion.

n	Indicators	Scale	Score	Weight	Max Score
28	land consumption // natural surfaces	not present	0	3.5	35
		present	10		
29	amount of natural surfaces (m^2)	low [< 10,000 m^2]	0	2	20
		medium [10,000-100,000 m^2]	5		
		high [>100,000 m^2]	10		

Information

The criterion of Information refers to the systematic provision of knowledge about a transport system to assist in the planning and execution of a journey (CEN/TC 320. 2002). Specifically, we focused on the second-level parameter of Information about accessibility (*Table 19*). The benchmark indicators refer to the presence of signs on the access to LPT services and infrastructure (stops), pedestrian paths, and cycleways, as well as tactile information for the access to different transport systems and about disabled-friendly interchanges (Diana et al., 2016).

Table 19. Indicators for information criterion.

n	Indicators	Scale	Score	Weight	Max Score
30	signs about the access to the LPT infrastructure/service	not present	0	3.5	35
		present	10		
31	signs about the access to the pedestrian infrastructure	not present	0	3	30
		present	10		
32	signs about the access to the cycling infrastructure	not present	0	3	30
		present	10		
33	provision of tactile information	not present	0	3.5	35
		present	10		
34	information about disabled friendly interchanges	not present	0	3.5	35
		present	10		

Safety and Security

The criterion of Safety and Security refers to the sense of personal protection experienced by travellers and transport users (CEN/TC 320, 2002). Included in this category are the first-level parameters of Emergency Management and Freedom from accident and crime. In this work, we focus on the latter two, represented by the second-level parameters of Traffic flow and Lighting, respectively, and we consider the indicators of motorized and non-motorized flow, and the presence of lighting (*Table 20*). These types of indicators refer to the presence of a working lighting system at the public transport stop, along the footpath and cycleway; and to the quantity of motorized vehicle (heavy and light), considering as low flow < 300 vehicles per hour (vph), medium 300 - 1,000 vph, and high >1,000 vph (D'Orso & Migliore, 2019). These data can be obtained from official sources (as in the case of this paper), if available, or through field surveys. Lighting and traffic flow are among the most widely recognized safety parameters in the literature (Bellizzi et al.,2019; Vallejo-Borda et al., 2020; Ujjwal, 2021; Du & Huang, 2022).

Table 20. Indicators for safety and security criterion.

n	Indicators	Scale	Score	Weight	Max Score
35	motorized flow (number of motorized vehicle per hour)	high [>1,000 vph]	0	3.5	35
		medium [300-1,000 vph]	5		
		low [<300 vph]	10		
36	working lighting system at the LPT stops	not present	0	3	30
		present	10		
37	working lighting system along the pedestrian infrastructure	not present	0	3	30
		present	10		
38	working lighting system along the cycling infrastructure	not present	0	3	30
		present	10		

Time

The criterion of Time refers to the aspects of time relevant to the planning and execution of journeys (CEN/TC 320, 2002). Included in this category are the first-level parameters of Adherence to schedule/Reliability, and Length of trip time. We focus on the latter one represented by the second-level parameters of At Origin/boarding and alighting/Destination points (*Table 21*). The related indicators refer to the time (min/h) spent waiting for the LPT and looking for a shared micro-mobility means of transport, as well as the time spent to reach de destination by LPT, walking, and cycling.

Table 21. Indicators for time criterion.

n	Indicators	Scale	Score	Weight	Max Score
39	min/h spent waiting the LPT	>5 min	0	1.5	15
		0-5 min	10		
40	min/h spent looking for a shared micro-mobility means of transport	>5 min	0	1	10
		0-5 min	10		
41	min/h spent to reach the destination with LPT	>20 min	0	1.5	15
		10-20 min	5		
		<10	10		
42	min/h spent to reach the destination by walking	>20 min	0	2.5	25
		10-20 min	5		
		<10	10		
43	min/h spent to reach the destination by cycling	>20 min	0	2	20
		10-20 min	5		
		<10	10		

Table 22 reports the results of the quality assessment conducted across the nine routes, summarising the score of each criterion and its corresponding indicators. For each criterion, the maximum attainable score is obtained by summing the scores of all indicators associated with that criterion. When a criterion is composed of multiple parameters (for example, Comfort includes both Cleanliness and Amenities), the table presents the maximum score for each parameter as well as the overall aggregated score for the entire criterion.

Table 22. Scores and assessment of indicators for the eight quality criteria.

Indicators	**Routes**								
	R1	R2	R3	R4	R5	R6	R7	R8	R9
Accessibility / Internal interface / Entrance/exits									
ramps at LPT stops	0	35	0	35	35	35	35	35	35
ramps to get on and off to the pedestrian infrastructure	35	35	35	35	35	35	35	35	35
ramps to enter and exit to the cycling infrastructure	35	35	35	35	35	35	35	35	35
Accessibility total score	**70**	**105**	**70**	**105**	**105**	**105**	**105**	**105**	**105**
Availability / Network / Area covered									
LPT service/infrastructure	0	35	0	35	35	35	35	35	35
pedestrian infrastructure	35	35	35	35	35	35	35	35	35
cycling infrastructure	35	35	35	35	35	35	35	35	35
bike-sharing service	35	35	35	35	35	35	35	35	35
scooters-sharing service	0	0	0	0	0	0	0	0	0
continuity of the pedestrian infrastructure	35	35	35	35	35	35	35	35	35
continuity of the cycle infrastructure	0	0	0	35	35	35	35	35	35
Availability total score	**140**	**175**	**140**	**210**	**210**	**210**	**210**	**210**	**210**
Comfort / Ambient conditions / Cleanliness									
bins at the LPT stops	0	10	0	10	10	10	10	10	10
bins along the pedestrian infrastructure	10	10	10	10	10	10	10	10	10
trash/broken glasses at the LPT stops	10	0	10	10	0	10	10	10	10
trash/broken glasses along the pedestrian infrastructure	0	0	0	10	10	10	10	10	10
trash/broken glasses along the cycling infrastructure	10	10	0	10	10	10	10	10	10
illegal graffiti at the LPT stops	10	10	10	10	10	10	10	10	10
illegal graffiti along the pedestrian infrastructure	0	10	10	10	0	10	0	0	0
illegal graffiti along the along the cycling infrastructure	0	10	10	10	0	10	0	0	0
Cleanliness score	**40**	**60**	**50**	**80**	**50**	**80**	**60**	**60**	**60**
Comfort / Complementary facilities / Amenities									
seating facility at the LPT stops	0	20	20	20	20	20	20	20	20
seating facility along the pedestrian infrastructure	15	15	15	15	15	15	15	15	15
public toilet	10	10	10	10	10	10	10	10	10
bicycle racks at interchanges	10	0	0	10	10	10	10	0	0
bicycle racks along the route	5	5	5	5	5	5	5	5	5
drinking water facility	10	10	10	10	10	10	10	10	10
number of drinking water facility	2.5	2.5	2.5	5	5	2.5	5	5	2.5
Amenities score	**52.5**	**62.5**	**62.5**	**75**	**75**	**72.5**	**75**	**65**	**62.5**
Comfort total score	**92.5**	**122.5**	**112.5**	**155**	**125**	**152.5**	**135**	**125**	**122.5**
Customer care / Assistance / Customer orientation									
customer service	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5
Customer care / Commitment / Innovation & initiatives									
access to a hotline	20	20	20	20	20	20	20	20	20
Customer care total score	**37.5**	**37.5**	**37.5**	**37.5**	**37.5**	**37.5**	**37.5**	**37.5**	**37.5**
Environmental impact /Pollution and health / Natural resources									
land consumption // natural surfaces	35	0	35	35	35	35	35	35	35
amount of natural surfaces (m^2)	12.5	0	25	25	12.5	25	25	25	25
Environmental impact total score	**47.5**	**0**	**60**	**60**	**47.5**	**60**	**60**	**60**	**60**
Information / General information / Info. about accessibility									
signs about the access to the LPT infrastructure/service	35	35	35	35	35	35	35	35	35
signs about the access to the pedestrian infrastructure	30	30	30	30	30	30	30	30	30
signs about the access to the cycling infrastructure	30	30	30	30	30	30	30	30	30
provision of tactile information	35	35	35	35	35	35	35	35	35
information about disabled friendly interchanges	35	35	35	35	35	35	35	35	35
Information total score	**165**	**165**	**165**	**165**	**165**	**165**	**165**	**165**	**165**
Safety & Security / Freedom from accidents / Traffic flow									
motorized flow (number of motorized vehicle per hour)	0	0	0	0	0	0	0	0	0

Safety & Security / Freedom from crime / Lighting									
working lighting system at the LPT stops	0	30	0	30	30	30	30	30	30
working lighting system along the pedestrian infrastructure	30	30	30	30	30	30	30	30	30
working lighting system along the cycling infrastructure	30	30	30	30	30	30	30	30	30
Freedom from crime score	**60**	**90**	**60**	**90**	**90**	**90**	**90**	**90**	**90**
Safety & Security total score	**60**	**90**	**60**	**90**	**90**	**90**	**90**	**90**	**90**
Time / Length of trip time / At Origin/boarding and alighting/Destination points									
min/h spent waiting the LPT	0	15	0	15	15	0	0	15	15
min/h spent looking for a shared micro-mobility means of transport	10	10	0	10	10	0	10	10	0
Time / Length of trip time / Transfer time									
min/h spent to reach the destination with LPT	0	10	0	7.5	15	0	15	7.5	15
min/h spent to reach the destination by walking	0	0	0	0	12.5	0	0	12.5	25
min/h spent to reach the destination by cycling	20	20	10	20	20	20	10	20	20
Time total score	**30**	**55**	**10**	**52.5**	**72.5**	**20**	**35**	**65**	**75**
Overall score of the path	**505**	**600**	**515**	**695**	**675**	**660**	**657.5**	**677.5**	**685**

Figure 28 summarizes the aggregated total scores, computed by eqn. 18, for each criterion of the quality measure and illustrates the comparative scores of the nine routes (R nº). The values of each indicator can be seen in non-aggregated form in *Table 22.* The table reports the results of the quality assessment conducted across the nine routes, summarising the score of each criterion and its corresponding indicators. For each criterion, the maximum attainable score is obtained by summing the scores of all indicators associated with that criterion. When a criterion is composed of multiple parameters (for example, Comfort includes both Cleanliness and Amenities), the table presents the maximum score for each parameter as well as the overall aggregated score for the entire criterion.

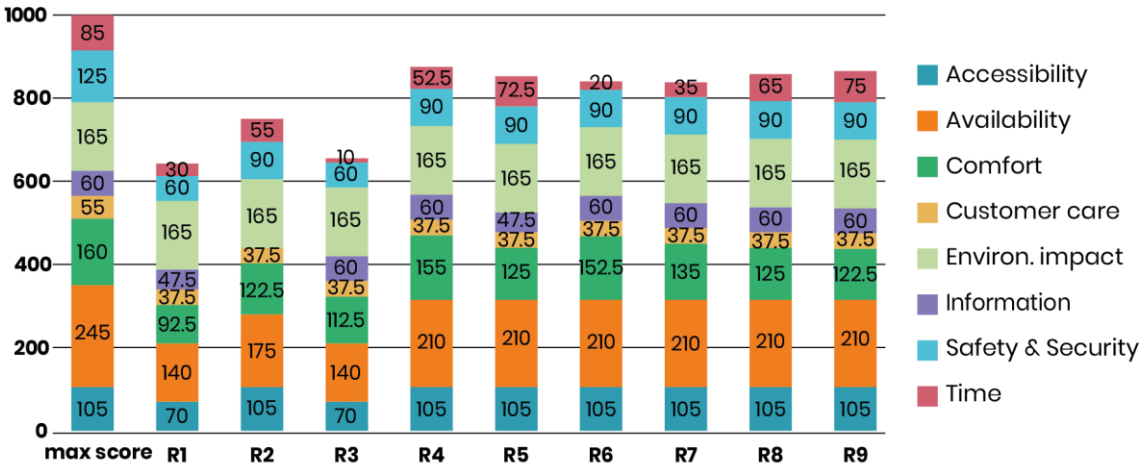


Figure 28. Comparison of maximum possible weighted score for each quality criteria with scores of the nine routes (R) in Melbourne, Australia.

The analysis of the nine urban routes enabled a comprehensive assessment of the strengths and weaknesses characterising the multimodal transport environment in the selected area. The evaluation was carried out using the full set of indicators associated with the eight quality criteria - Accessibility, Availability, Comfort, Customer Care, Environmental Impact, Information, Safety & Security, and Tim - thereby offering an integrated picture of the overall quality of the system.

Overall, the results show a heterogeneous quality profile across the routes. Routes 4, 9, and 8 achieved the highest aggregated weighted scores, while Routes 1 and 3 displayed the lowest. Among the quality criteria, Information consistently achieved the maximum score across all routes. This reflects the widespread presence of signage, tactile paving, and accessibility information across the network, highlighting a consolidated practice of ensuring legibility and wayfinding.

By contrast, Availability never reached the maximum score on any route, mainly due to the systematic absence of shared e-scooter services. In Routes 1 and 3, the absence of public transport services along the walking path also contributed to lower quality. Similarly, Environmental Impact scored relatively highly in most routes, but Routes 1, 2, and 5 showed lower values due to the limited or absent presence of natural areas.

Other criteria - Comfort, Customer Care, Safety & Security, and Time - did not achieve maximum values on any route. Comfort is particularly affected by issues such as litter, broken glass, and graffiti along several paths. Customer Care showed consistent but moderate values due to the predominance of virtual over physical customer support across the network. Safety & Security was penalised mainly by the high flow of motorised vehicles. Time-related indicators showed that cycling systematically offered the most efficient travel times, whereas walking was convenient only on selected routes, and public transport waiting times were occasionally excessive (e.g., in Routes 6 and 7).

Taken together, the indicators captured the main criticalities affecting the multimodal walking-cycling-micromobility-LPT interface. From an operational perspective, several targeted improvements emerge. Enhancing shared micromobility availability could strengthen last-mile accessibility, consistent with evidence from the literature on the value of e-scooter systems (Baek et al., 2020). Comfort could be improved through better street maintenance, while Customer Care would benefit from a greater presence of physical support points at interchanges. Safety could be reinforced by interventions aimed at reducing motorised traffic, in combination with policies discouraging private car use. Finally, time efficiency could be improved by increasing LPT service frequencies in selected routes.

The set of indicators introduced in this study confirms their usefulness for monitoring multimodal transport environments. Their quantitative nature, combined with the ease of data acquisition, makes the framework suitable for supporting planning and design decisions, particularly those aimed at enhancing multimodal interchange and improving the perceived quality of the urban mobility offer. The results from this preliminary

assessment provide a valuable basis for refining the methodology and validating it indifferent settings, such as the Verona case study presented in the following sections.

5.2.2 Case study II: Verona, Italy

Verona, located in the northeastern part of Italy, has a population of approximately 260,000 inhabitants and features a compact historic city centre surrounded by mixed-use peripheries. The city presents an interesting case for the study of urban mobility due to its established multimodal transport infrastructure and historic urban fabric, which influence accessibility and the distribution of transport services.

The urban transport system in Verona comprises a network of local roads, bus services connecting the city and its outskirts, regional railway stations, and a growing array of shared mobility options, including bike-sharing and scooter sharing. While the city demonstrates a reasonable integration of multiple transport modes, challenges remain in terms of quality.

Verona was selected as a case study because it exemplifies a medium-sized Italian city with a complex and multimodal mobility system. Its characteristics, including the coexistence of traditional LPT, emerging shared mobility solutions, and a constrained historic centre, make it an ideal context for evaluating the quality of urban multimodal transport.

The main objective of this study is to assess the quality of multimodal transport in Verona by analysing eight quality criteria of availability, accessibility, comfort, customer service, environmental impact, information, safety and security, and time. The assessment focuses on a set of indicators reflecting these criteria and representative of the KQPs selected in phase 1 of the overall methodology, allowing for a systematic evaluation of multimodal mobility quality at the urban scale.

The assessment of multimodal transport quality in Verona has been carried out through a dual approach. On the one hand, a subjective perspective was incorporated through a survey distributed among citizens, which provided insights into expected and perceived quality, based on user experience (Customer satisfaction technique); on the other hand, an objective evaluation was conducted by means of field-based and desk-based analysis (Mystery Client technique). This combination allows for a more comprehensive understanding of multimodal transport quality in the city.

In the Verona case study, the Customer Satisfaction (CS) survey was conducted prior to the Mystery Client (MC) assessment, as the selection of routes for the MC measurements was guided by the survey responses. Specifically, routes were chosen for the objective assessment phase based on the origin and destination neighbourhoods of respondents, focusing on areas with the highest number of survey responses.

5.2.2.1 Customer Satisfaction (CS)

An online survey[3] was developed and distributed among residents of Verona, Italy, to assess citizens' perceptions of the quality of the urban transport system. The survey targeted daily travel patterns between respondents' residential neighborhoods (points of origin) and the main destinations for regular activities such as work, education, sports, or leisure.

The survey specifically focused on non-motorized mobility (walking and cycling), LPT (urban buses), and electric micro-mobility solutions (e-scooters and e-bikes). Participants were asked to evaluate ten quality parameters previously identified through the systematic literature review. For each parameter, two independent scores were collected on a 1-10 Likert scale: one representing the perceived importance of the parameter and the other reflecting the level of satisfaction derived from personal experience. This dual-rating structure enables the comparison between expected and perceived quality, thereby identifying potential quality gaps and priorities for improvement.

The survey was administered anonymously and disseminated through institutional mailing lists, community networks, and local social media channels. Data collection was carried out during June and July 2025, and the responses were analyzed in aggregate form to ensure confidentiality and statistical reliability. The findings of this survey are used to complement the results of the Mystery Client analysis, providing an integrated understanding of subjective and objective quality dimensions within the Verona case study.

Sample characteristics and Origin-Destination information

A total of 165 valid responses were collected. As the survey was openly accessible, the number of individuals reached is unknown, and a response rate could not be calculated.

The survey participants were asked to indicate their main occupational status. The majority of respondents (67%) identified themselves as workers, followed by retirees (27%). A smaller proportion consisted of students (3%) and individuals who selected other occupations (3%) (*Figure 29*).

While the sample does not fully reflect the demographic structure of Verona's population, it provides valuable insights into the perspectives of two key groups. This distribution suggests that the sample predominantly reflects the perspective of the working population, with a significant share of retired individuals also represented. Although students and other categories were less represented, the diversity of occupational

[3] The full version of the survey is available at: https://docs.google.com/forms/d/e/1FAIpQLSdeqIHKzRIhP9bOn4klAp6bBNB8-5-tNrRhNYs8Axp0Ia3FJw/viewform?usp=sf_link

backgrounds contributes to a broader understanding of urban mobility experiences across different demographic groups.

Respondents were also asked to indicate their place of residence. The majority of participants (84%) reported living within the municipal boundaries of the City of Verona, while the remaining 16% reside in other municipalities within the Province of Verona.

This spatial distribution ensures that the survey primarily captures the perceptions and mobility experiences of urban residents, while also incorporating the perspective of commuters or peri-urban dwellers who may interact with the city's transport system from outside its core.

When disaggregating residential location by occupational status, the distribution appears relatively homogeneous: workers (83% city, 17% province) and retirees (84% city, 16% province) show nearly identical patterns, while all students reported living within the city, and respondents in the 'other' category were similarly concentrated in the urban area (80% city, 20% province).

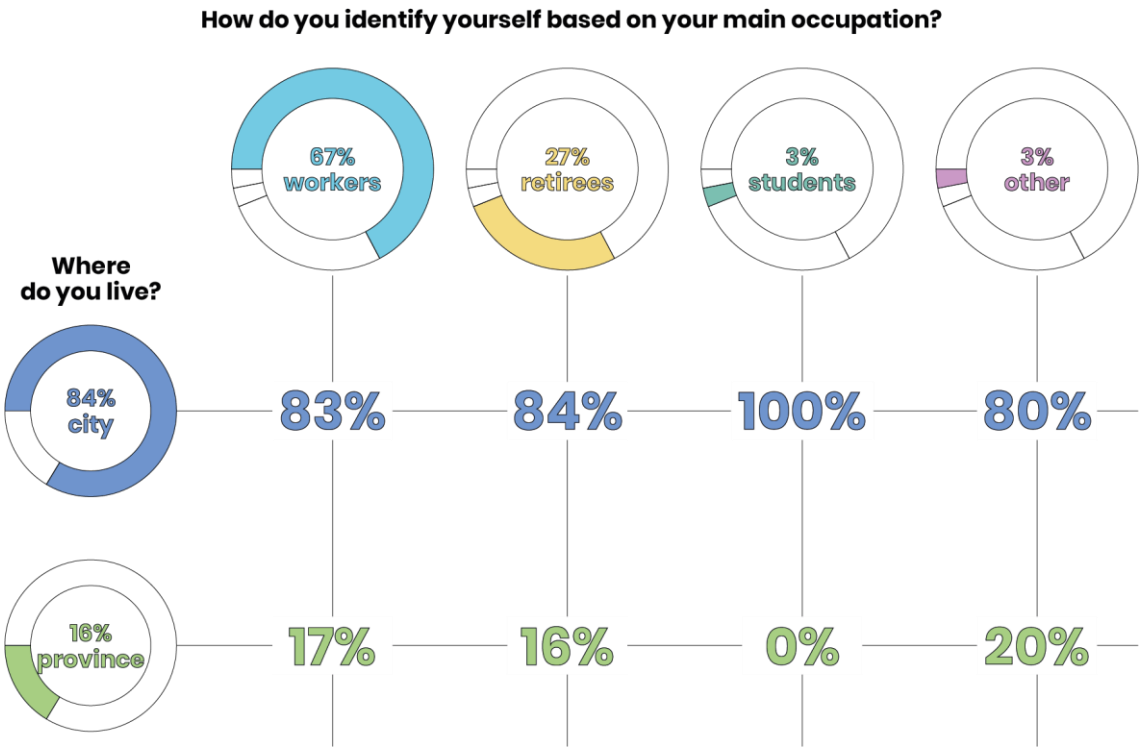


Figure 29. Distribution of survey respondents by occupational status and area of residence (city vs. province).

Among the 165 total respondents, 84% indicated that they reside within the City of Verona, while the remaining 16% live in other municipalities within the Province of Verona. Only participants residing in the city were asked to specify their neighborhood of residence, as the survey aimed to investigate mobility patterns originating from specific urban districts.

The spatial distribution of the 138 urban residents shows a broad territorial representation across the municipal area. The most frequently reported neighborhoods were Borgo Milano (including the localities of Chievo and Saval) and Valdonega, each with 16% of respondents who come from the city. These were followed by Borgo Trento (9%), Borgo Venezia (including Borgo Trieste) (7%), and the historic city center (Città Antica) (6%). Other neighborhoods such as Ponte Crencano, Veronetta, San Zeno, Cadidavid, San Massimo, and San Michele were also represented with moderate frequencies (4%-5%).

Less populated responses came from Borgo Roma, Montorio, Avesa, Cittadella, Quinzano, and Santa Lucia (3%), while neighborhoods such as Parona, Porto San Pancrazio, Golosine, Mizzole, Quinto, and Santa Maria in Stelle were each reported at 1%-2%.

Figure 30 illustrates the percentage-based spatial distribution of survey respondents across neighbourhoods of origin.

This wide intra-urban distribution of respondents supports a more comprehensive understanding of how transport quality is perceived across different types of neighborhoods, from central and historic districts to more suburban areas.

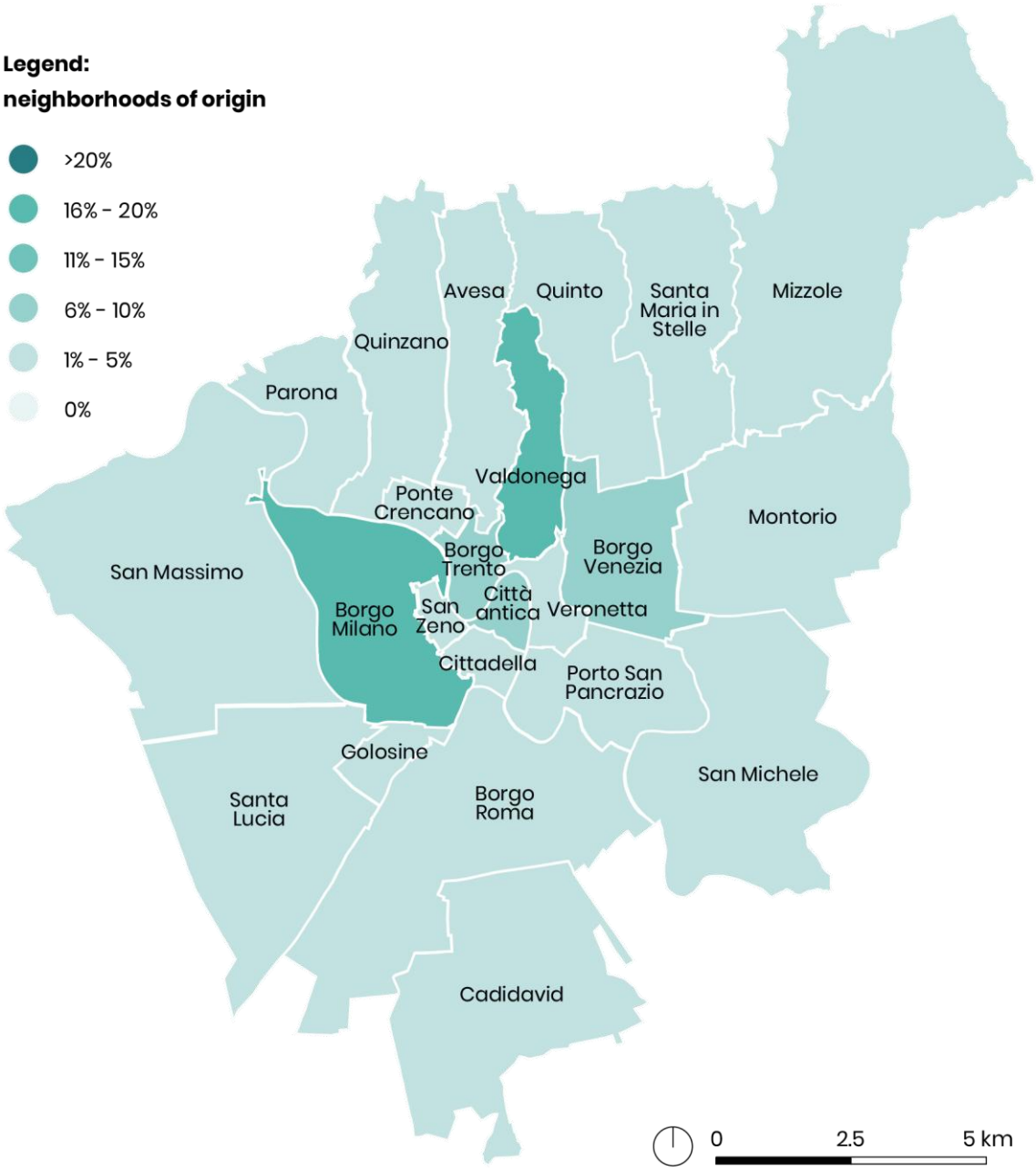


Figure 30. Spatial distribution of survey respondents from the city by neighbourhood of origin (O).

Out of the 138 respondents residing within the City of Verona, 83% reported that their regular travel destination is also located within the urban area. The remaining 17% respondents indicated destinations located outside the city and were therefore excluded from this specific question.

The responses show that the historic city center (Città Antica) is by far the most frequented destination, with 32% of city-based commuters identifying it as their primary activity hub. This is followed by Cittadella, the neighborhood encompassing Verona Porta Nuova railway station, with 18% of respondents, highlighting the strategic role of this area as a multimodal transport and service node. In fact, the survey explicitly instructed participants who regularly reach destinations outside the city but use the train as their main mode of travel to select the Cittadella district as their destination, since it includes Verona Porta

Nuova station. This ensured that the analysis accurately reflected the station area's importance as a gateway to extra-urban mobility.

Other frequently cited destination neighborhoods include Borgo Roma (10%), Borgo Trento (9%), and Borgo Milano (8%). Areas such as Veronetta (5%), Borgo Venezia (4%), and Golosine and Valdonega (3%) also appear as relevant zones for daily activities.

A few respondents reported destinations in neighborhoods of Ponte Crencano, San Michele, and San Zeno (2%), and Avesa, Cadidavid, Porto San Pancrazio, and Santa Lucia were each cited by the 1%.

Figure 31 illustrates the percentage-based spatial distribution of survey respondents across neighbourhoods of destination.

This distribution underscores the role of central and semi-central districts as major attractors of daily mobility flows, while also reflecting the diversity of activity centers throughout the urban fabric.

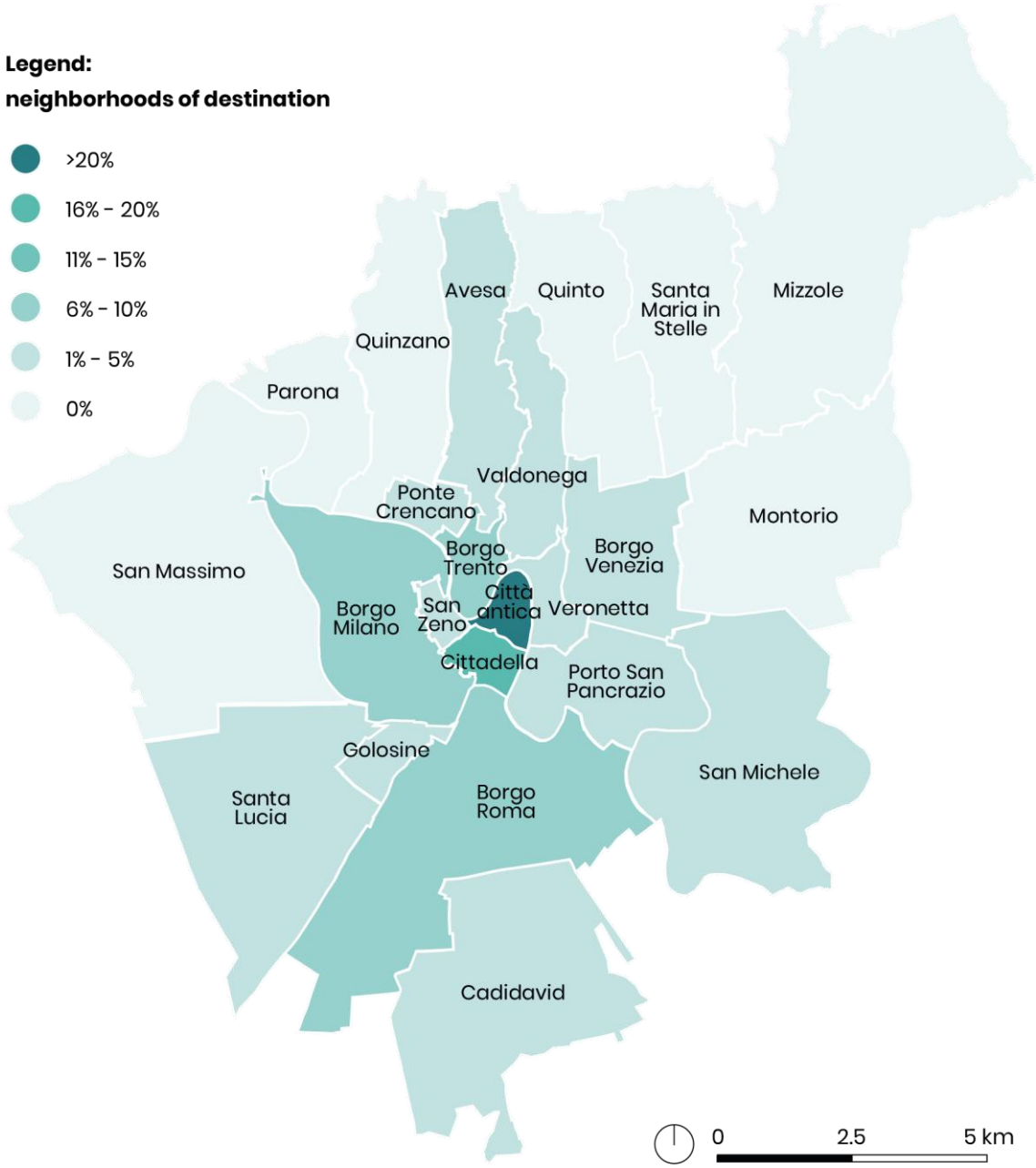


Figure 31. Spatial distribution of survey respondents from the city by neighbourhood of destination (D).

Respondents were asked to indicate the main modes of transport they use on a daily or frequent basis, including multimodal combinations. The responses reveal a significant diversity in travel behaviors, with a strong presence of both non-motorized and private motorized mobility (*Figure 32*).

The most commonly reported mode was private bicycle or electric bicycle (20%), highlighting the relevance of cycling infrastructure and the potential of non-motorized mobility within the city. This was followed by combinations involving private vehicles and non-motorized modes, such as private car or motorcycle + bicycle (15%), private car or motorcycle + walking (14%), and private car alone (12%).

These data indicate a widespread reliance on private vehicles, often integrated with more sustainable modes. LPT was present in various multimodal configurations, particularly in combination with walking (9%) or with private vehicles (7%), while motorcycles as a standalone mode were reported by 6%. Some participants also indicated combinations involving walking and cycling (6%) or walking alone (5%), further reinforcing the importance of active mobility options. Less frequently mentioned combinations included LPT + bicycle (3%), bicycle + electric scooter (1%), and walking + micromobility services such as bike-sharing or e-scooter sharing (1%). Only one respondent selected “Other”, while isolated combinations involving private vehicles and sharing services (bike-sharing or e-scooter sharing) were each reported once.

These results suggest a multimodal mobility landscape where private car use remains prominent but is frequently integrated with active modes. At the same time, the data points to a notable role for cycling and walking in everyday travel choices, as well as a still marginal adoption of shared micromobility services.

%	transport modes	sustainability	multimodality
20%	Own bicycle or electric bicycle	high	low
15%	Private car or motorcycle + Own bicycle or e-bicycle	medium	medium
14%	Private car or motorcycle + Walking	medium	medium
12%	Private car	low	low
9%	Local Public Transport + Walking	high	medium
7%	Private car or motorcycle + Local Public Transport		
6%	Walking + Own bicycle or electric bicycle	high	medium
6%	Motorcycle	medium	medium
5%	Walking	low	low
3%	Local Public Transport + Own bicycle or electric bicycle	high	low
1%	Walking + Bike-sharing	high	medium
1%	Walking + Shared scooter	high	medium
1%	Other	high	medium
1%	Private car or motorcycle + Bike-sharing	medium	medium
1%	Private car or motorcycle + Shared scooter	medium	medium
		low level medium level high level	low level medium level high level

Figure 32. Distribution of transport modes most frequently used by respondents living in the city.

In addition to the 138 urban residents, the survey included 24 respondents residing in municipalities outside the City of Verona, within the surrounding provincial area. These participants were asked to indicate their main mode of transport used to reach the city on a daily or frequent basis.

The results show a strong reliance on private motorized transport, with 75% of respondents indicating the private car as their primary mode of access. A smaller number reported using a motorcycle (13%), while only 4% stated that use an extra-urban bus service. The 8% selected “Other”, but did not specify the mode (*Figure 33*).

For respondents residing in the provincial area and commuting to the city, travel behaviour appears significantly less sustainable and entirely non-multimodal. The vast majority (96%) rely exclusively on private motorized vehicles, with 75% traveling by private car and a further 13% by motorcycle. Only a very small minority (4%) reported the use of interurban bus services, representing the sole instance of a fully sustainable choice within this subgroup.

From the perspective of multimodality, no respondent indicated the use of more than one mode of transport, resulting in a pattern of exclusively unimodal trips. This sharply contrasts with the urban residents, where moderate multimodality was more widespread.

Overall, the results suggest that provincial commuters are heavily dependent on private motorized mobility, with limited adoption of LPT and no integration of different modes. This highlights a structural challenge for the promotion of sustainable and multimodal mobility options.

These findings highlight the challenges of promoting sustainable and collective mobility options in suburban and peri-urban contexts, where the private vehicle remains the dominant and often most practical choice due to infrastructural or service limitations. The extremely limited use of public or shared transport services suggests the need for targeted interventions to improve connectivity and accessibility from the province to the city, particularly in the context of commuting.

%	transport modes	sustainability	multimodality
75%	Private car	○	○
13%	Motorcycle	○	○
8%	Other		
4%	Intercity bus	○	○
		○ low level ◐ medium level ● high level	○ low level ● medium level ● high level

Figure 33. Distribution of transport modes most frequently used by respondents living in the province.

Even respondents residing in the Province of Verona were also asked to indicate their final destination within the city, i.e., the neighbourhood where they most frequently travel for work, study, errands, or leisure purposes.

The most frequently cited destination was Città Antica (29%), followed by Borgo Roma and Borgo Trento (21%), which together account for over 70% of the reported destinations. These areas represent key urban hubs characterized by a high concentration of services, educational institutions, healthcare facilities, and commercial activities.

Other destinations included Borgo Milano (8%), Cittadella (8%) - which, as previously mentioned, contains the main railway station and serves as a multimodal node - and more peripheral or residential districts such as Cadidavid, Ponte Crencano, and Quinzano (8% each).

Figure 34 illustrates the percentage distribution of destination neighbourhoods in the city, as reported by respondents originating from the provincial area.

Although limited in number, these responses show that provincial commuters are primarily oriented toward central and semi-central neighbourhoods, reinforcing the centripetal nature of urban mobility flows toward areas of high functional density.

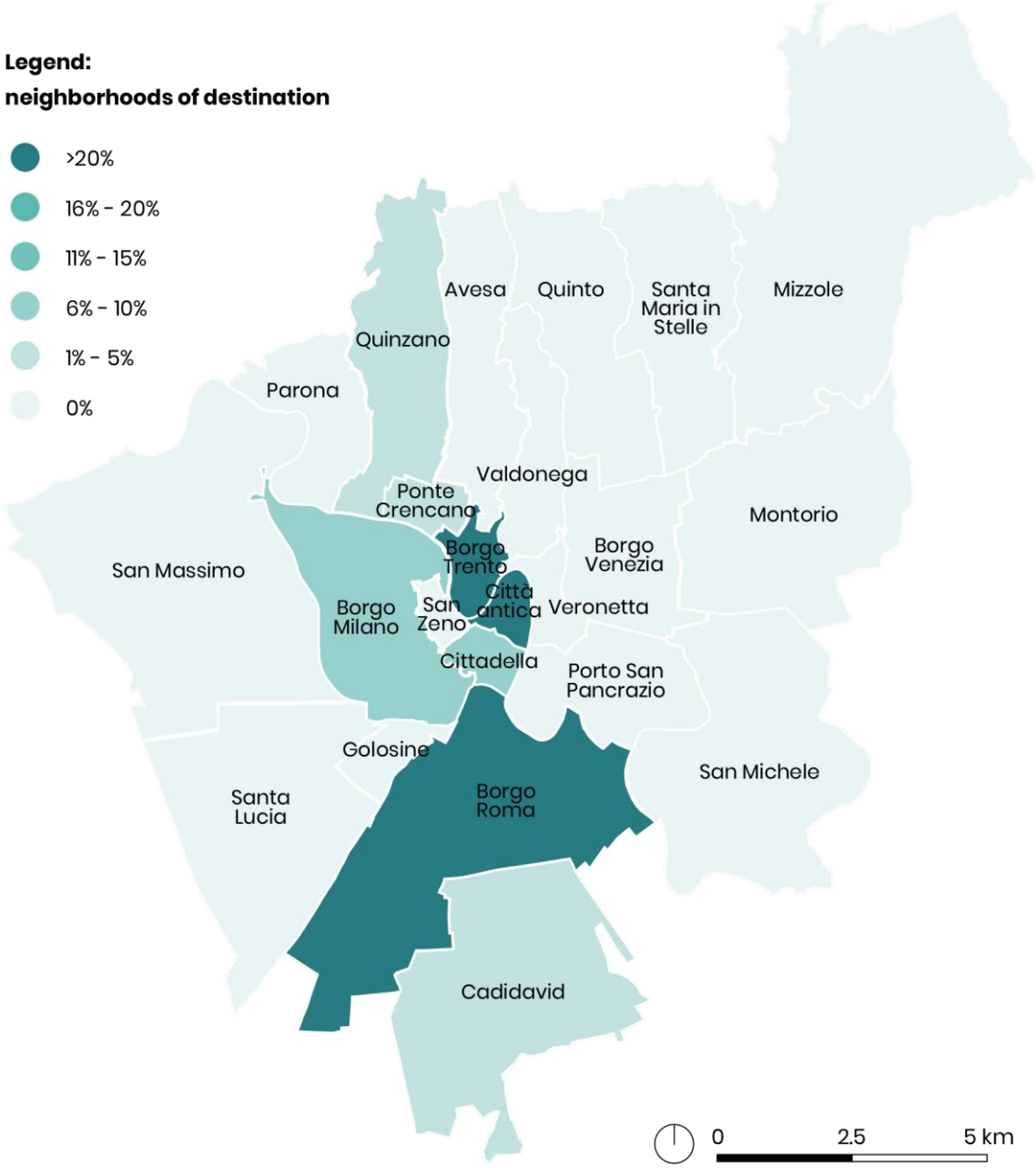


Figure 34. Spatial distribution of destination neighbourhood (D) in the city, as reported by respondents originating from the provincial area.

By combining the information on reported trip origins and destinations, an O-D matrix was constructed to illustrate the spatial distribution of travel flows between neighbourhoods (*Figure 35* and *Figure 36*). The matrix, which considers only the responses provided by residents living within the municipal boundaries, highlights the main intra-urban connections and reveals the neighbourhoods that most frequently serve as points of departure and arrival. This representation provides a comprehensive overview of the mobility patterns within the city. Furthermore, it allows for the identification of areas that play a central role in shaping daily commuting dynamics. To support interpretation, the

results are presented both as a numerical O-D matrix and as a cartographic visualization, where the main flows between neighbourhoods are spatially highlighted.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1	Città antica	4	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	5 %
2	Cittadella	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3 %
3	San Zeno	2	1	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5 %
4	Veronetta	2	0	0	3	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6 %
5	Borgo Trento	4	2	0	0	4	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	11 %
6	Valdonega	4	8	0	0	2	2	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	17 %
7	B.go Venezia/B.go Trieste	4	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6 %
8	Porto San Pancrazio	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2 %
9	Borgo Roma	0	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4 %
10	Santa Lucia	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3 %
11	B.go Milano / Chievo / Saval	7	2	1	0	0	0	0	0	2	0	2	0	0	0	0	0	0	0	1	0	0	0	0	16 %
12	Golosine	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1 %
13	Ponte Crencano	0	1	0	0	1	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	4 %
14	Avesa	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	3 %
15	Quinto	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1 %
16	Santa Maria in Stelle	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 %
17	Mizzole	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0 %
18	Montorio	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1 %
19	San Michele	2	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	4 %
20	Cadidavid	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3 %
21	San Massimo	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3 %
22	Parona	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0 %
23	Quinzano	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	3 %
		32 %	18 %	2 %	5 %	9 %	3 %	4 %	1 %	10 %	1 %	8 %	3 %	2 %	1 %	0 %	0 %	0 %	0 %	2 %	1 %	0 %	0 %	0 %	100 %

Figure 35. O-D matrix based on survey respondents.

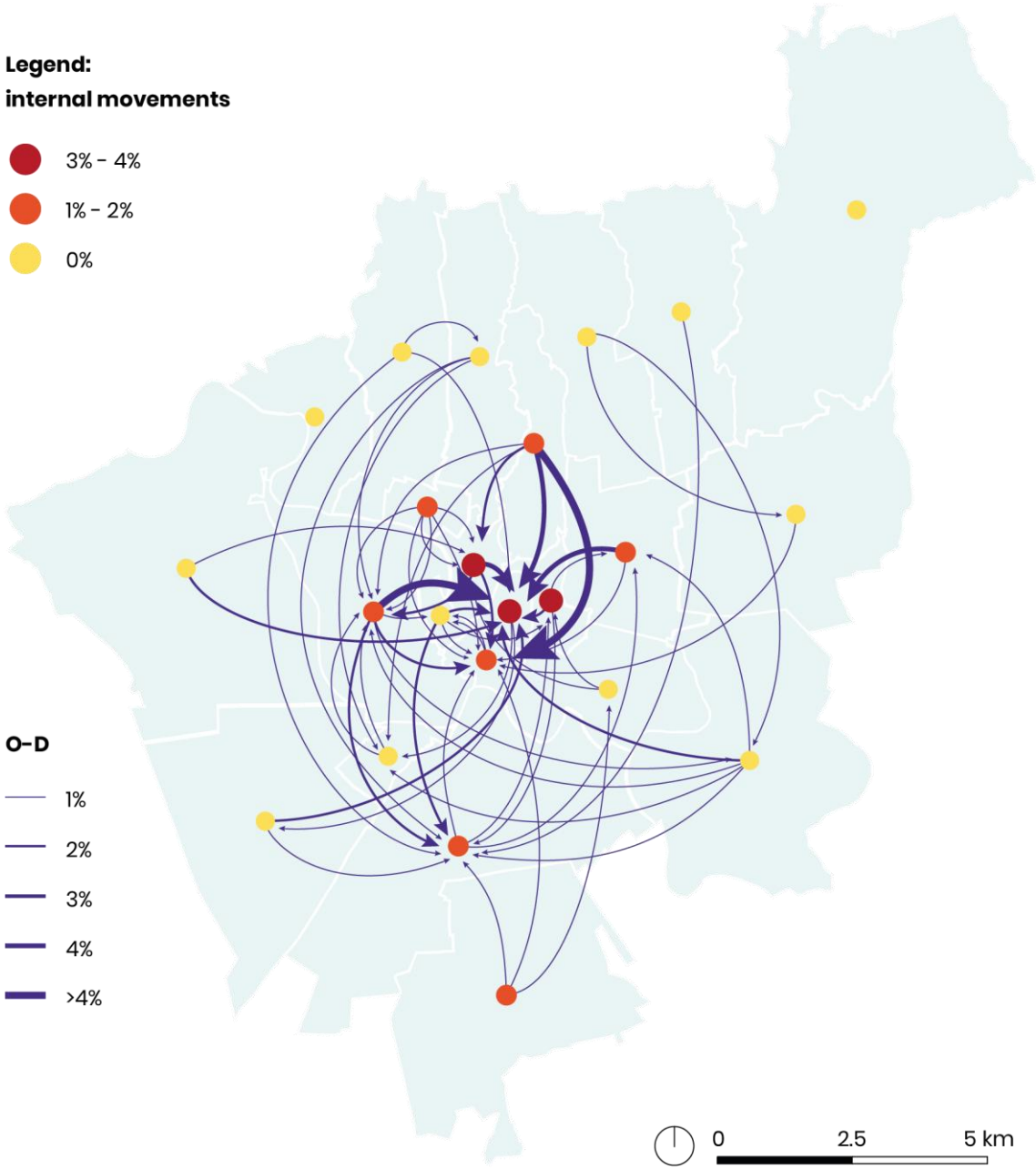


Figure 36. Spatial distribution of travel flows between neighbourhoods of origin (O) and destination (D) in the city.

The combined analysis of the O-D matrix and the cartographic visualization reveals that the highest concentrations of movements originate from central neighborhoods such as Città Antica and Cittadella, while significant flows also involve Borgo Trento, Valdonega, Borgo Roma, and Borgo Milano/Chievo/Saval. These areas emerge as both major origins and destinations, reflecting their functional importance within the city. The map further highlights the strong connections between central and semi-central neighborhoods, as well as the relatively weaker flows from peripheral areas. Overall, the visual evidence confirms that daily mobility patterns are strongly polarized towards the historical core and

its immediate surroundings, making these connections particularly relevant for subsequent quality assessments.

The O-D matrix was used to identify the neighbourhood pairs with the highest concentration of responses, revealing the most frequently travelled connections within the sample. These O-D pairs provide the spatial framework for the subsequent analysis, ensuring that both subjective (expected and perceived quality) and objective (delivered quality) assessments focus on areas that are most representative of users' actual mobility patterns.

Four O-D pairs emerged as the most significant and were therefore selected for further investigation: Borgo Trento–Città Antica, Borgo Milano–Città Antica, Valdonega–Cittadella, and Borgo Venezia–Città Antica. These corridors not only include the central urban area, which is the most frequent destination among respondents, but also capture flows from different parts of the city: north (Valdonega), east (Borgo Venezia), and west (Borgo Milano). This spatial diversity ensures a balanced representation of travel behaviour and supports a robust evaluation of quality across varied urban contexts.

Quality measurement based on expected quality (EQ) and perceived quality (PQ)

The CS survey provides a set of scores reflecting users' expectations (EQ) and perceived quality (PQ) for each Key Quality Parameter across different modes of transport in the selected O-D pairs. The results allow a detailed overview of how respondents evaluate the relevance of each parameter and the quality they experience when travelling.

Transport modes vary from one O-D to another (*Figure 37, Figure 38, Figure 39, Figure 40*), as they depend on how respondents to the survey travel most frequently.

Across all four O-D pairs, the CS dataset offers a comprehensive representation of:

- multiple modes of transport selected by respondents;
- variation in expectations and perceived service levels;
- evaluation of all eight quality criteria (Availability, Accessibility, Time, Comfort, Customer Care, Environmental Impact, Information, Safety & Security) through their ten associated KQPs.

The quality measurement results constitute the empirical basis for the subsequent Quality Management phase, where EQ and PQ will be compared to assess the magnitude and direction of quality gaps for each mode and parameter.



Figure 37. Average values of expected (EQ) and perceived (PQ) quality for users travelling along O-D 1 (Borgo Trento – Città Antica), disaggregated by the transport modes reported in the survey.

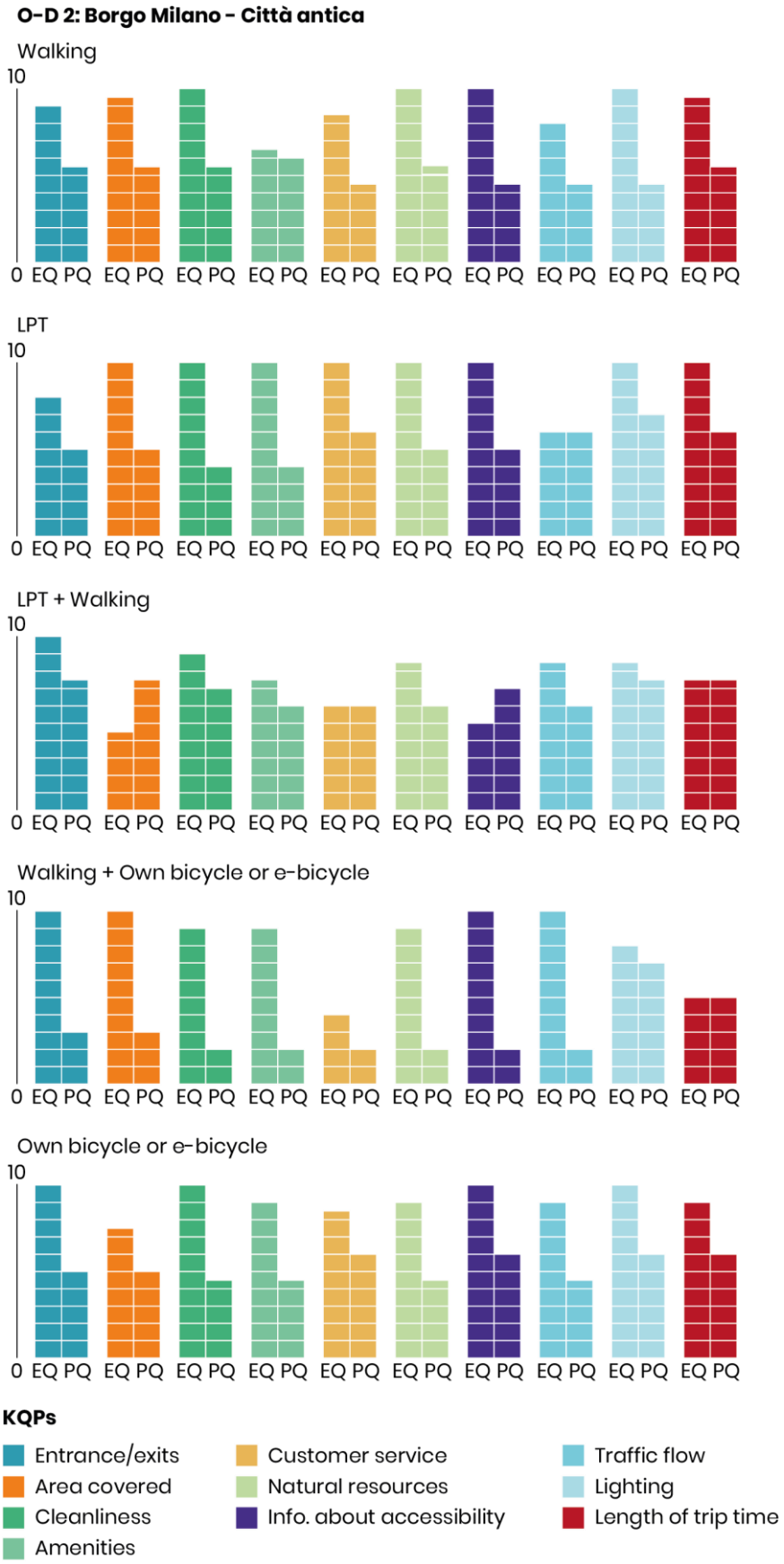


Figure 38. Average values of expected (EQ) and perceived (PQ) quality for users travelling along O-D 2 (Borgo Milano – Città Antica), disaggregated by the transport modes reported in the survey.

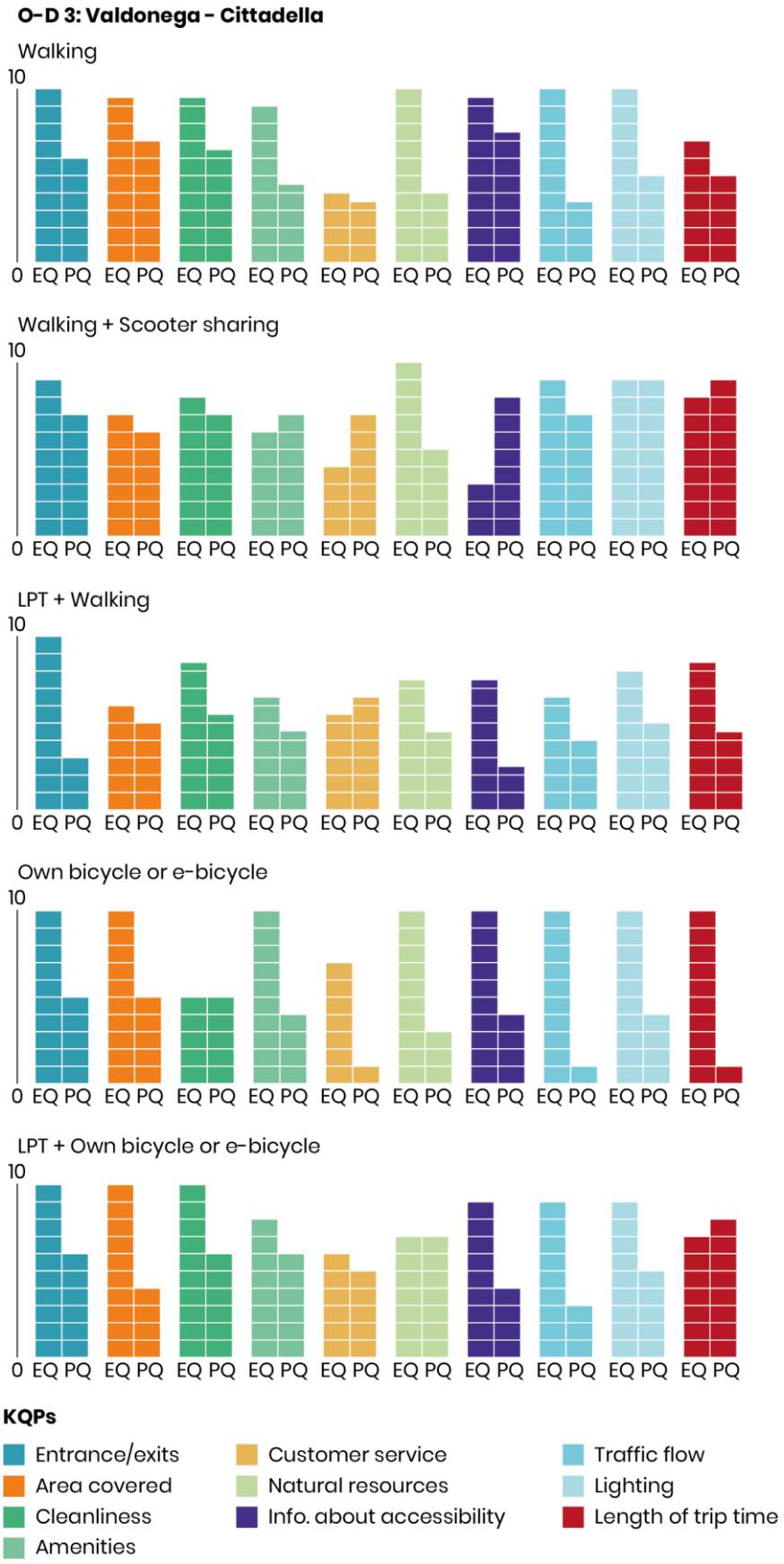


Figure 39. Average values of expected (EQ) and perceived (PQ) quality for users travelling along O-D 3 (Valdonega – Cittdella), disaggregated by the transport modes reported in the survey.

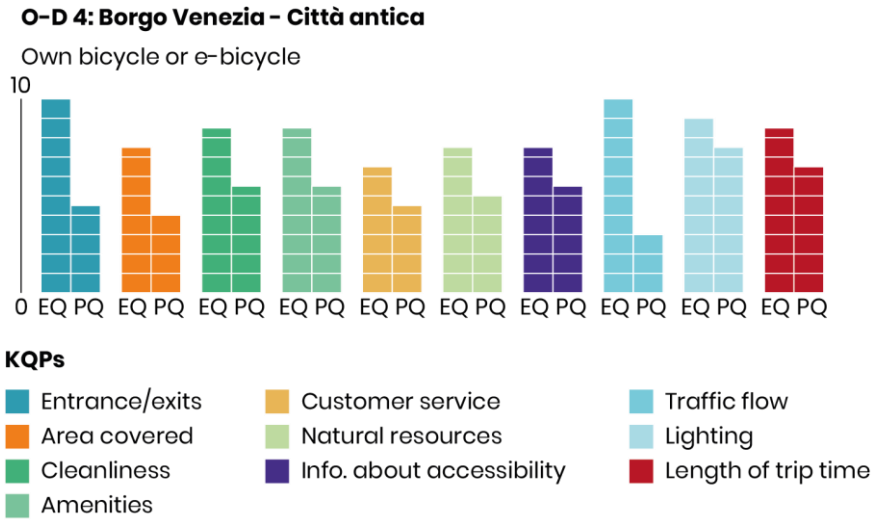


Figure 40. Average values of expected (EQ) and perceived (PQ) quality for users travelling along O-D 4 (Borgo Venezia – Città Antica), disaggregated by the transport modes reported in the survey.

This measurement approach allows the analysis to be performed at both a disaggregated and an aggregated level while explicitly capturing the multimodal nature of urban travel. At the disaggregated level, EQ and PQ scores are examined separately for each individual transport mode as well as for the combinations of modes reported by respondents (e.g., walking + LPT, bicycle + e-scooter). This enables the identification of quality patterns specific to single-mode and multimodal travel behaviours. At the aggregated level (*Figure 41*) mode-specific scores - whether referring to single or combined modes - are averaged to obtain a single EQ and PQ value for each KQP or criterion within every O-D pair. This dual perspective makes it possible to analyse both the diversity of multimodal user experiences and the overall quality profile of each corridor.

	O-D 1		O-D 2		O-D 3		O-D 4	
	EQ	PQ	EQ	PQ	EQ	PQ	EQ	PQ
Accessibility	8.13	5.57	8.30	5.20	8.30	5.40	7.50	4.00
Availability	7.75	7.00	9.40	5.20	9.80	5.40	10.00	4.25
Comfort	8.81	5.38	9.10	4.65	8.05	5.60	8.75	5.38
Customer care	3.75	3.63	7.40	4.90	5.27	4.57	6.75	5.40
Environmental impact	8.00	5.13	9.50	4.70	8.87	4.67	7.50	5.00
Information	10.00	4.75	9.00	4.90	7.83	5.23	7.75	5.50
Safety & Security	9.50	5.75	8.80	5.50	9.07	4.85	9.50	5.13
Time	5.63	8.50	8.20	6.00	8.33	6.53	8.25	6.75

Figure 41. Average values of expected (EQ) and perceived (PQ) quality for users travelling along the four selected O-Ds, aggregated by quality criteria.

Although *Figure 41* presents only the aggregated EQ and PQ scores for each O-D pair and each quality criterion, these values represent the basis for the subsequent step of the measurement. According to the Eqns 19-20, the expected and perceived quality ratings are then converted into the corresponding percentages of users in conformity with the designed quality (PUI) and with the delivered quality (PUS). These conformity measures provide a more operational and policy-relevant understanding of the quality levels, allowing the assessment to move from a simple comparison of scores to an evaluation of how many users actually experience (or expect) acceptable quality conditions. The calculation and interpretation of PUI and PUS are presented below (*Figure 42*).

	O-D 1		O-D 2		O-D 3		O-D 4	
	PUI	PUS	PUI	PUS	PUI	PUS	PUI	PUS
Accessibility	80	40	75	38	63	50	75	25
Availability	60	80	100	38	100	38	100	25
Comfort	80	60	81	50	75	50	75	50
Customer care	80	60	100	38	75	75	75	50
Environmental impact	0	20	50	50	0	50	50	25
Information	100	60	88	50	75	38	50	25
Safety & Security	100	25	82	64	88	31	100	38
Time	40	100	75	75	88	50	75	75

Figure 42. Percentage of users for whom each KQP is important (PUI) and percentage of users satisfied (PUS) across the four selected O-D pairs, aggregated by quality criteria.

5.2.2.2 Mystery Client (MC)

The Mystery Client (MC) assessment was conducted on the same four O-D pairs identified through the Customer Satisfaction (CS) survey. Aligning the objective evaluation with the subjective responses ensures methodological coherence across the quality measurement phase, allowing the delivered quality to be assessed precisely where users most frequently travel. The selected O-D pairs - Borgo Trento–Città Antica, Borgo Milano–Città Antica, Valdonega–Cittadella, and Borgo Venezia–Città Antica (*Figure 43)* - thus constitute the core spatial framework for the MC fieldwork.

For each O-D pair, all segments composing the identified paths were travelled and assessed in real operating conditions using the structured MC protocol described in Chapter 4. This parallel structure enables a consistent comparison between expected, perceived, and delivered quality, and supports an integrated interpretation of multimodal mobility quality across Verona.

Furthermore, for each O-D pair, the delivered-quality assessment was conducted using the segmentation and path-selection procedure described in Chapter 4 (Appendix D, *Table D.2*).

The paths identified for each O-D pair differ in length and structure. Each O-D includes two or three alternative paths, selected either because they represent the main connecting routes between the corresponding neighbourhoods or because they were explicitly mentioned by respondents in the survey through the optional question asking them to indicate the streets typically used when travelling between that specific origin and destination. For each path, its total length and the number of segments into which it was subdivided for the Mystery Client assessment are reported in *Table 23*.

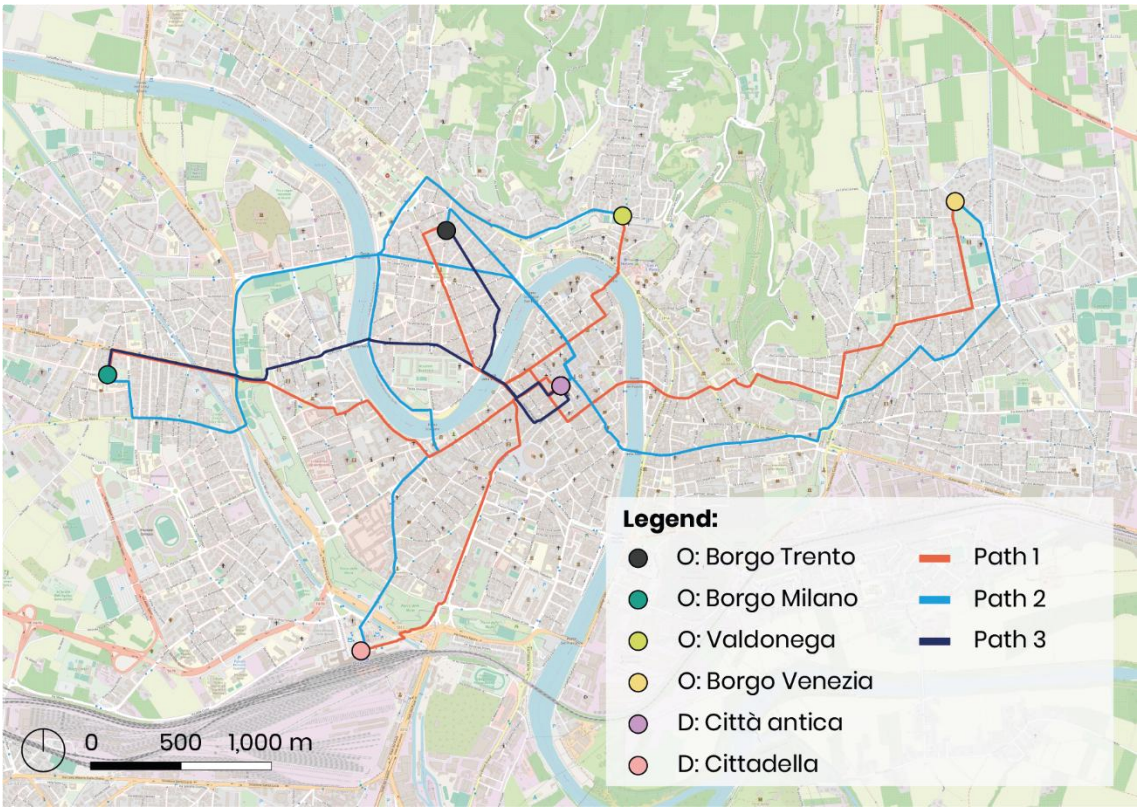


Figure 43. Spatial representation of the paths connecting the four selected Origin-Destination (O-D) pairs.

Table 23. Characterisation of the four selected O-Ds.

O-D	Origin (O)	Destination (D)	N° of paths	Paths km	N° of segments
O-D 1	Borgo Trento	Città antica	3	1,600	4
				1,440	3
				1,472	3
O-D 2	Borgo Milano	Città antica	3	3,012	7
				4,140	9
				3,074	7
O-D 3	Valdonega	Cittadella	2	3,343	7
				4,314	9
O-D 4	Borgo Venezia	Città antica	2	3,588	8
				3,754	8
O-D tot			**10**	**29,737**	**65**

Following the identification of the paths selected for the MC, the next step involves defining the set of indicators used to measure delivered quality across all segments, paths, and O-D pairs. The indicators adopted in this application correspond to those introduced in the first empirical test, but their weighting structure differs. As described in Chapter 4, the weights used here were derived through a multi-criteria comparison process, in which several researchers independently performed pairwise comparisons between indicators belonging to the same quality criterion (Appendix D, *Table D.1*). The resulting weights were averaged to obtain a single synthetic value for each indicator, whose stability was assessed through the analysis of standard deviation and coefficient of variation. The low dispersion values confirmed the robustness of the final weighting scheme, ensuring that indicator relevance is consistently reflected in the computation of delivered quality.

Table 24 reports, for each quality criterion, the indicators included in the assessment, their measurement scale, and their assigned weight. This provides the operational basis for the subsequent calculation of delivered quality using the hierarchical procedure outlined in the methodology.

Table24. Quality criteria, parameters, and indicators used in the Mystery Client assessment, including their final weights derived through the multi-criteria evaluation procedure.

Parameter I	Parameter II	Indicator	Scale	Score	Weigh	Max score
Accessibility						**1,000**
Internal interface	Entrance/exits	Ramps at LPT stops	not present	0	0.285	285
			present	1		
		Ramps to get on and off to the pedestrian infrastructure	not present	0	0.263	263
			present	1		
		Ramps to enter and exit to the cycling infrastructure	not present	0	0.452	452
			present	1		
Availability						**1,000**
Network	Area covered	LPT service/infrastructure	not present	0	0.172	172
			present	1		
		Pedestrian infrastructure	not present	0	0.203	203
			present	1		
		Cycling infrastructure	not present	0	0.138	138
			present	1		
		Bike-sharing service	not present	0	0.081	81
			present	1		
		Scooter-sharing service	not present	0	0.056	56
			present	1		
		Continuity of the pedestrian infrastructure	interrupted	0	0.195	195
			continuous	1		
		Continuity of the cycling infrastructure	interrupted	0	0.155	155
			continuous	1		
Comfort						**2,000**
Ambient condition	Cleanliness	Bins at the LPT stops	not present	0	0.103	103
			present	1		
		Bins along the pedestrian infrastructure	not present	0	0.101	101
			present	1		
		Trash/broken glasses at the LPT stops	not present	0	0.205	205
			present	1		
			not present	0	0.206	206

		Trash/broken glasses along the pedestrian infrastructure	present	1		
		Trash/broken glasses along the cycling infrastructure	not present	0	0.206	206
			present	1		
		Illegal graffiti at the LPT stops	not present	0	0.069	69
			present	1		
		Illegal graffiti along the pedestrian infrastructure	not present	0	0.053	53
			present	1		
		Illegal graffiti along the cycling infrastructure	not present	0	0.055	55
			present	1		
Complementary facilities	Amenities	Seating facility at the LPT stops	not present	0	0.216	216
			present	1		
		Seating facility along the pedestrian infrastructure	not present	0	0.111	111
			present	1		
		Public toilet	not present	0	0.141	141
			present	1		
		Bicycle racks at interchanges	not present	0	0.190	190
			present	1		
		Bicycle racks along the route	not present	0	0.096	96
			present	1		
		Drinking water facility	not present	0	0.130	130
			present	1		
		Number of drinking water facility	not present	0	0.117	117
			present	1		
Customer care						**2,000**
Assistance	Customer orientation	Customer service	not present	0	1	1,000
			present	1		
		Hotline	not present	0	1	1,000
			present	1		
Environmental impact						**1,000**
Pollution and health	Natural resources	land use // natural surfaces	not present	0	0.375	375
			present	1		
		amount of natural surfaces (m^2)	low [<10.000 m^2]	0	0.625	625
			medium/high [>10.000m^2]	1		
Information						**1,000**
General information	Info. about accessibility	Signs about the access to the LPT infrastructure/service	not present	0	0.239	239
			present	1		
		Signs about the access to the pedestrian infrastructure	not present	0	0.166	166
			present	1		
		Signs about the access to the cycling infrastructure	not present	0	0.142	142
			present	1		
		Tactile information at nodal points (intersections, major crossings)	not present	0	0.226	226
			present	1		
		Information about disabled friendly interchanges	not present	0	0.226	226
			present	1		
Safety & Security						**2,000**
Freedom form accident	Traffic flow	motorized flow (number of motorized vehicle per hour)	medium/high [>300 vph]	0	1	1,000
			low [<300 vph]	1		
Freedom form crime	Lighting	working lighting system at the LPT stops	not present	0	0.239	239
			present	1		
		working lighting system along the pedestrian infrastructure	not present	0	0.427	427
			present	1		
		working lighting system along the cycling infrastructure	not present	0	0.333	333
			present	1		
Time						**2,000**

Length of trip time	At O/boarding & alighting/D points	min/h spent waiting the LPT	> 5 min	0	0.688	688
			0-5 min	1		
		min/h spent looking for a shared micro-mobility means of transport	> 5 min	0	0.313	313
			0-5 min	1		
	Transfer time	min/h spent to reach the destination with the LPT	>15 min	0	0.528	528
			<15 min	1		
		min/h spent to reach the destination by walking	>20 min	0	0.205	205
			<20 min	1		
		min/h spent to reach the destination by cycling	>15 min	0	0.267	267
			<15 min	1		
All criteria						**12,000**

Quality measurement based on delivered quality

The Mystery Client (MC) technique yields a set of delivered quality (DQ) scores for each KQP across the four selected O-D pairs. These scores result from the weighted aggregation of the indicators listed in *Table 24* and reflect the actual conditions experienced along the observed routes.

According to the Eqns 13-17, delivered quality for the four selected O-D pairs is computed through the hierarchical aggregation of segment-, path-, and OD-level scores. However, because the MC dataset does not include user-level observations, the Percentage of Users in Conformity (PUC) is estimated using the proportional proxy described in the methodological note of Chapter 4. Specifically, the score obtained for each KQP is normalised with respect to its maximum attainable value, producing a percentage (PUC_MC) that expresses the extent to which the designed quality standard is effectively met along the observed routes. This procedure ensures full consistency with the theoretical framework while adapting the computation to the nature of MC-derived data.

Across all O-D pairs, the MC dataset provides a complete operational measurement of the eight quality criteria - Availability, Accessibility, Time, Comfort, Customer Care, Environmental Impact, Information, and Safety & Security - through the final KQP-level and criterion-level aggregations.

The delivered quality results presented in the following *Figure 44* constitute, together with the expected (EQ) and perceived (PQ) quality outcomes derived from the CS technique, the empirical foundation for the subsequent Quality Management phase. In this final stage of the assessment, Expected Quality, Perceived Quality, Delivered Quality, and Designed Quality will be jointly compared to quantify the magnitude and direction of quality gaps at both the individual KQPs and overarching criteria levels.

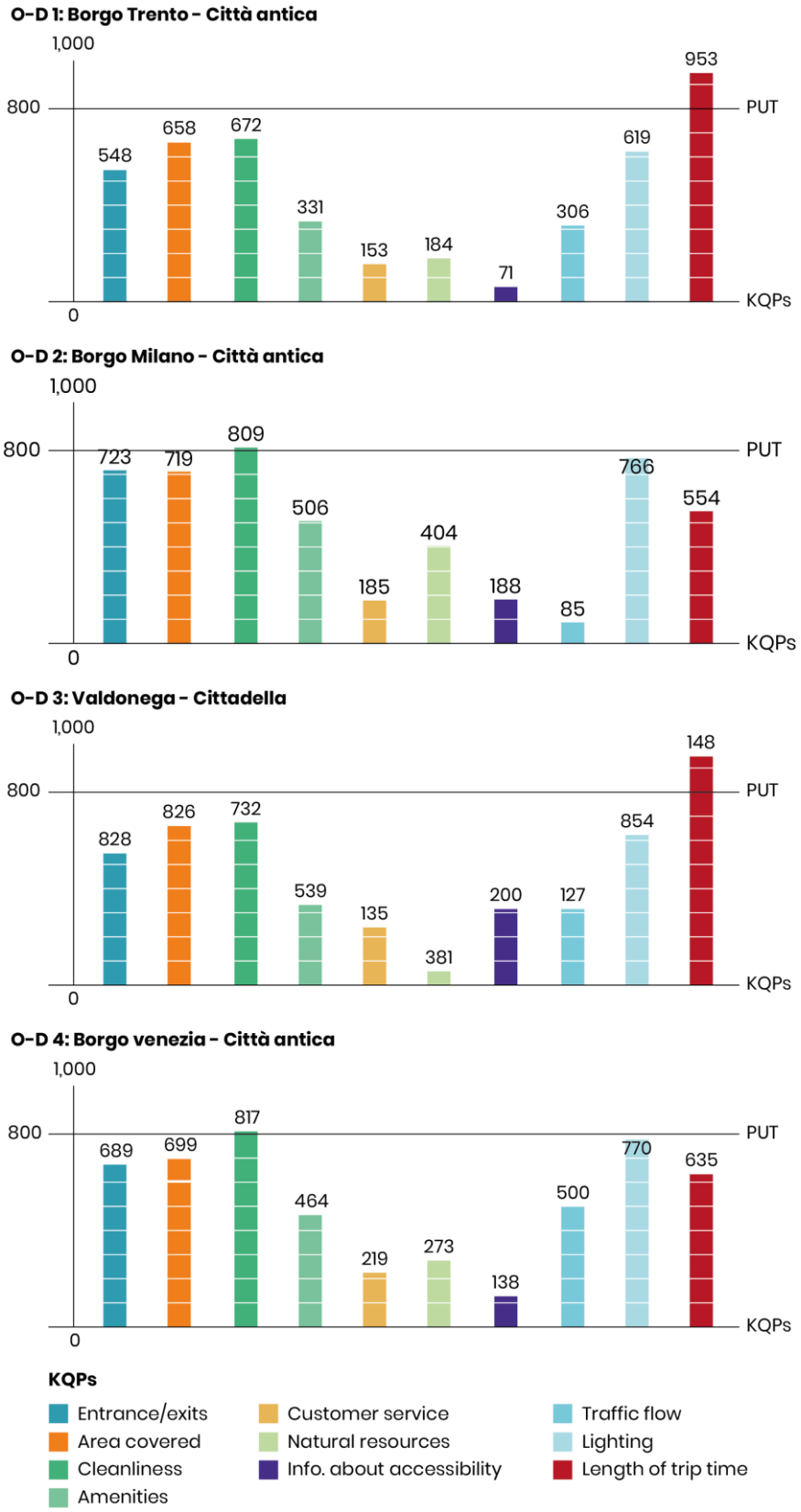


Figure 44. Average values of delivered quality objectively measured along the four O-Ds through selected KQPs.

Across the four O-D pairs, the delivered quality scores reveal a heterogeneous quality of the multimodal network. Some KQPs consistently exhibit higher values, such as entrances/exits, area covered, and cleanliness, indicating that these aspects of the system tend to maintain relatively robust conditions across corridors. Other parameters show more variable behaviour. Indicators such as natural resources, amenities, and traffic flow display notable fluctuations between O-Ds, reflecting spatial differences in exposure to motorised traffic.

Certain KQPs, including information about accessibility and customer orientation, register comparatively lower scores, suggesting that these aspects are less uniformly guaranteed along the measured segments. Time-related attributes also differ substantially across O-Ds.

Overall, *Figure 44* illustrates the distinctive quality profile of each corridor, emphasising that the delivered quality is not homogeneous across the city but depends on both the physical characteristics of the routes and the distribution of services and infrastructure along them.

At the aggregated level (*Figure 45),* the KQPs scores are averaged to obtain a single value for delivered quality for each criterion within every O-D pair.

	O-D 1 DQ	O-D 2 DQ	O-D 3 DQ	O-D 4 DQ
Accessibility	548	723	828	689
Availability	658	719	826	699
Comfort	502	657	636	640
Customer care	153	185	135	219
Environmental impact	184	404	381	273
Information	71	188	200	138
Safety & Security	462	426	490	635
Time	953	554	148	184

Figure 45. Average values of designed quality (DQ) along the four selected O-Ds, aggregated by quality criteria.

Although *Figure 45* presents only the aggregated delivered quality scores for each O-D pair and each KQP, these values constitute the foundation for the next step of the measurement process. Consistent with the methodological framework outlined in Chapter 4, the delivered quality values obtained through the MC assessment are converted into the Percentage of Users in Conformity with the designed quality (PUC_MC).

This transformation enables the analysis to move beyond raw indicator scores and toward an operational evaluation of how many users would receive a level of service consistent with the predefined standards. The PUC therefore provides a more actionable measure of delivered quality, allowing the assessment to quantify the share of users who are expected to experience acceptable conditions along each O-D corridor.

The calculation and interpretation of PUC_MC values are presented below (*Figure 46),* forming the link between the objective field measurements and the subsequent Quality Management phase.

In this case study, a single observation was made for each segment. This choice was motivated by the nature of most of the indicators assessed, which refer to infrastructure or physical characteristics that are stable over time and unlikely to vary on a daily basis (e.g., availability of seating, availability of cycling infrastructure, continuity of pavements).

Consequently, the quality provided was calculated following the hierarchical sequence of equations (13)–(17). However, unlike the general formulation, only one observation was made per segment, rather than multiple observations over time.

Furthermore, due to the absence of automatic or manual tools for counting users, the number of users was set at one for each segment. In this context, the Mystery Client acted simultaneously as an evaluator and a user. Consequently, each segment-level evaluation represents a single observation unit, rather than an estimate derived from a population of users.

	O-D 1 PUC_MC	O-D 2 PUC_MC	O-D 3 PUC_MC	O-D 4 PUC_MC
Accessibility	55	72	83	69
Availability	66	72	83	70
Comfort	50	66	64	64
Customer care	15	19	13	22
Environmental impact	18	40	3	27
Information	7	19	20	14
Safety & Security	46	43	49	64
Time	95	55	15	18

Figure 46. Percentage of Users in Conformity (PUC_MC) with the designed quality across the four selected O-D pairs, aggregated by quality criteria.

5.3 Quality management in multimodal transport systems

The Quality Management phase integrates the results of the Customer Satisfaction (CS) and Mystery Client (MC) techniques to evaluate how the mobility system performs relative to its intended level of quality. This phase is operationalised through the computation of four Quality Gaps (QG1-QG4), each reflecting a different misalignment between designed, delivered, expected, and perceived quality. By analysing these gaps across the four selected O-D pairs and the eight quality criteria, it becomes possible to identify the specific dimensions in which quality diverges from user needs or from the designed standards, as well as the drivers underlying such discrepancies. The gap patterns then provide the basis for defining targeted and criterion-specific improvement strategies.

5.3.1 Quality Gaps Analysis

The following sections interpret the gap matrices produced for the four O-D pairs. Although absolute values vary across criteria and O-Ds, consistent patterns emerge, revealing structural issues in both service provision and user perception.

Before examining gaps, the four quality dimensions are summarised for each criterion and for each of the four selected O-D pairs (*Table 25*).

These values constitute the basis for gap computation:

- PUT: designed quality threshold (80% across all criteria).
- PUC: % of users receiving quality measured objectively through MC.
- PUI: % of users whose expectations meet the designed quality.
- PUS: % of users perceiving the delivered quality as acceptable.

Table 25. Direct comparison of subjective (PUI, PUS) and objective (PUC) measures for each quality criterion and O-D.

CRITERIA	O-D 1				O-D 2				O-D 3				O-D 4			
	PUT	PUC	PUI	PUS	PUT	PUC	PUI	PUS	PUT	PUC	PUI	PUS	PUT	PUC	PUI	PUS
Accessibility	80	55	80	40	80	72	75	38	80	83	63	50	80	69	75	25
Availability	80	66	60	80	80	72	100	38	80	83	100	38	80	70	100	25
Comfort	80	50	80	60	80	66	81	50	80	64	75	50	80	64	75	50
Customer care	80	15	80	60	80	19	100	38	80	13	75	75	80	22	75	50
Environmental impact	80	18	0	20	80	40	50	50	80	38	0	50	80	27	50	25
Information	80	7	100	60	80	19	88	50	80	20	75	38	80	14	50	25
Safety & Security	80	46	100	25	80	43	82	64	80	49	88	31	80	64	100	38
Time	80	95	40	100	80	55	75	75	80	15	88	50	80	18	75	75

Quality gaps are computed using the four metrics according to the Eqns 21-24. QG1 = PUT – PUI represents the alignment between designed quality and users' expectations.

QG2 = PUC – PUT indicates whether delivered quality meets or exceeds the designed standard. QG3 = PUS – PUC indicates whether users perceive higher/lower quality than what is objectively delivered. Finally, QG4 = PUS – PUI represents the difference between expectations and perceived quality.

Table 26 shows the quality gaps for each criterion and O-D. Positive/negative values provide direct information on under-quality, over-quality, misalignment or perceptual bias.

Table 26. Quality Gaps for each quality criterion and O-D.

CRITERIA	O-D 1				O-D 2				O-D 3				O-D 4			
	QG 1	QG 2	QG 3	QG 4	QG 1	QG 2	QG 3	QG 4	QG 1	QG 2	QG 3	QG 4	QG 1	QG 2	QG 3	QG 4
Accessibility	0	-25	-15	-40	5	-8	-34	-37	17	3	-33	-13	5	-11	-44	-50
Availability	20	-14	14	20	-20	-8	-34	-62	-20	3	-45	-62	-20	-10	-45	-75
Comfort	0	-30	10	-20	-1	-14	-16	-31	5	-16	-14	-25	5	-16	-14	-25
Customer care	0	-65	45	-20	-20	-61	19	-62	5	-67	62	0	5	-58	28	-25
Environmental impact	80	-62	2	20	30	-40	10	0	80	-42	12	50	30	-53	-2	-25
Information	-20	-73	53	-40	-8	-61	31	-38	5	-60	18	-37	30	-66	11	-25
Safety & Security	-20	-34	-21	-75	-2	-37	21	-18	-8	-31	-18	-57	-20	-16	-26	-62
Time	40	15	5	60	5	-25	20	0	-8	-65	35	-38	5	-62	57	0

This subsection explains how the reader should interpret the table, without entering into case-specific interpretations (the actual interpretation will follow):

- QG1 close to zero suggests alignment between designed and expected quality.
- Negative QG2 indicates that delivered quality is below the designed quality threshold.
- Positive QG3 suggests that users perceive quality better than measured; negative indicates dissatisfaction beyond what measurement would predict.
- A high negative QG4 indicates a large expectation-experience mismatch.

This prepares the ground for the interpretation.

Interpretation of QG1 (PUT – PUI)

QG1 compares the importance attributed to each quality criterion by the system design (PUT) with the importance expressed by users (PUI). Positive values indicate that the system assigns more relevance to a criterion than users actually do, whereas negative values suggest that users perceive a criterion as more important than reflected in the design. Values close to zero denote good alignment between the two perspectives.

Overall, QG1 reveals a generally coherent relationship between designed and perceived importance, although the degree of alignment varies across criteria and O-D pairs. Accessibility presents consistently small positive gaps across all O-Ds, indicating that the

designed importance level is well matched with user expectations; the only slightly larger deviation in OD3 still remains within a range that suggests moderate rather than substantial misalignment. A similar pattern emerges for Comfort, where values close to zero confirm that design assumptions and user perceptions are largely aligned.

More variable behaviour is observed in criteria such as Availability and Customer Care. Availability shows a positive gap only in OD1, where the system overestimates its relevance, whereas in the other O-Ds the negative values indicate that users attribute greater importance to the criterion than foreseen by the design. Customer Care also presents mixed results: perfect alignment in OD1, underestimation of user expectations in OD2, and slight overestimation in OD3 and OD4. These fluctuations suggest that the perceived relevance of service-related aspects is highly context-dependent and may reflect different user profiles or expectations across city areas.

Environmental Impact stands out as the criterion with the largest positive gaps, particularly in OD1 and OD3. These large discrepancies indicate that the system consistently overestimates the relevance users attribute to environmental aspects. Although environmental impact remains a key component of sustainable mobility planning, the strong positive QG1 values suggest that users do not regard it as a primary determinant of service quality.

Information and Time exhibit heterogeneous results across the different corridors. For Information, negative gaps in the first two O-Ds indicate that users value the availability of guidance and signage more than assumed, whereas OD3 and especially OD4 show the opposite trend, suggesting that in those areas, information provision is less central to user expectations. Travel Time presents the strongest overestimation in OD1, while in OD3, the gap becomes negative, pointing to an underestimation of how important travel time is for those users. These variations reinforce the idea that the relevance of certain criteria is not homogeneous across the network.

A more consistent pattern emerges for Safety & Security, which shows negative values in all O-Ds, although with varying intensity. This finding indicates that users generally consider safety-related aspects as more important than reflected in the designed standards. The underestimation is particularly pronounced in OD1 and OD4, suggesting that concerns about traffic interactions, lighting, or perceived exposure are especially salient in these corridors.

Taken together, the QG1 results highlight that while several criteria - such as Accessibility and Comfort - are well calibrated in terms of designed versus perceived importance, others reveal systematic mismatches. In particular, Environmental Impact and, to a lesser extent, Travel Time tend to be overvalued in the design, whereas Safety & Security and Availability (in several O-Ds) tend to be underestimated by the system relative to user expectations.

Interpretation of QG2 (PUC – PUT)

QG2 measures the alignment between the designed quality standards (PUT) and the actual quality delivered as captured by the Mystery Client assessment (PUC). Negative values indicate that the service delivered does not reach the designed quality threshold, while positive values signal that the delivered quality exceeds the designed quality. Across all O-D pairs, QG2 reveals a widespread pattern of under-quality, with only a few isolated cases where the delivered quality approaches or slightly surpasses the predefined standards.

Accessibility displays moderate negative gaps in three of the four O-D pairs, with values ranging from –25 to –11, indicating that the infrastructure elements supporting accessible movement consistently fall short of their intended quality level. The only exception is OD3, where the small positive gap (+3) suggests that, for this specific corridor, the delivered accessibility marginally meets or slightly exceeds the expected standard.

A similar trend emerges for Availability, whose QG2 values remain negative across most O-Ds (from –14 to –10), again with a single marginally positive value in OD3 (+3). These results suggest that although public transport coverage, shared mobility presence, and walking/cycling continuity are relatively stronger in OD3, the rest of the network struggles to reach the expected levels of service provision. The consistency of slightly negative values points to broad, mild deficiencies rather than severe gaps.

The situation becomes more critical when examining Comfort, Customer Care, Environmental Impact, and Information. Comfort displays negative gaps across all O-Ds, ranging between –30 and –14, reflecting shortcomings in cleanliness, amenities, and the general user experience along the assessed corridors. The breadth and consistency of these gaps suggest that comfort-related aspects systematically fail to achieve the expected standard, revealing structural deficiencies in the pedestrian and multimodal environment.

Customer Care presents the most pronounced gaps in the entire dataset, with extremely high negative values across all O-Ds (between –58 and –67). These results reflect the near-absence of customer service elements in the measured environment. The consistently large negative values highlight a structural gap intrinsic to the service context rather than isolated shortcomings.

Environmental Impact also shows a uniform set of substantial negative values, ranging from –62 to –40. These values suggest that aspects related to land use integration of the urban surroundings consistently underperform relative to the desired standard.

Information is another criterion showing systematic and severe under-quality, with QG2 values between –73 and –60 across all O-Ds. This indicates that information about accessibility is consistently insufficient when compared to the designed expectations. Among all criteria, Information and Customer Care stand out as those where the gap between expected and delivered quality is structurally the most pronounced.

Safety & Security presents moderately negative values in all O-Ds except OD4, where the gap is less severe (–16 compared to values around –31 to –37 in the others). These results indicate general deficiencies in lighting, traffic interaction safety, or perceived exposure along the corridors. The variation across O-Ds suggests that safety conditions are not uniformly distributed in the network, with some environments offering relatively better, though still insufficient-quality.

Finally, the Time criterion displays a more heterogeneous pattern. OD1 records a positive QG2 value (+15), indicating that delivered travel times exceed the expected standard, likely due to favourable congestion conditions or relatively direct alignments. By contrast, OD2 and especially OD3 and OD4 show substantial negative values (–25, –65, –62), signalling that in most cases travel times exceed the thresholds considered acceptable in the design.

Across all criteria and O-D pairs, QG2 highlights widespread misalignment between expected and delivered quality, with most delivered values falling significantly below the designed standards. Although some criteria display mild under-quality and isolated cases of slight overachievement, the overall trend suggests structural deficiencies in several dimensions of the mobility environment, particularly Customer Care, Information, Environmental Impact, and Comfort. These findings emphasise the importance of targeted interventions aimed at narrowing the gap between planning expectations and the actual quality experienced on the ground.

Interpretation of QG3 (PUS – PUC)

QG3 measures the gap between the proportion of users who report being satisfied with the service (PUS) and the proportion who, according to the objective Mystery Client assessment, actually receive a service that conforms to the designed quality standard (PUC). Positive values indicate that users perceive the service more favourably than what is objectively delivered, whereas negative values indicate that users are less satisfied than the objectively measured quality would suggest.

Across the four O-D pairs, Accessibility displays consistently negative QG3 values (from –15 to –44), suggesting that users are systematically less satisfied with accessibility conditions than what the objective assessment would imply. The magnitude of the negative values indicates that accessibility is a dimension where subjective experience diverges substantially from objective compliance.

Availability shows a more mixed pattern. OD1 presents a positive gap (+15), suggesting that users perceive the availability of transport options as better than the objective assessment indicates. However, in the remaining three O-D pairs, the gap is strongly negative (between –34 and –45), revealing widespread dissatisfaction despite a moderate level of objectively measured provision. Comfort also presents moderate positive gaps in OD1 (+10), indicating that users sometimes perceive the comfort level to

be higher than its actual measured quality. In the remaining corridors, the values are consistently negative (–14 to –16), pointing to a general tendency for comfort conditions, such as cleanliness and amenities, to be perceived as less satisfactory than what the MC inspection registers.

Customer Care shows the highest positive QG3 values across the entire dataset (from +19 to +62), meaning that users consistently report higher satisfaction with customer care than what is objectively offered. The very large positive gaps suggest a systematic overestimation of available support services.

Environmental Impact exhibits small but generally positive QG3 values (between +2 and +12 for three O-Ds), indicating slightly higher user satisfaction than objective conditions would suggest. This may reflect the presence of greenery or pleasant streetscape elements that, although not sufficient to meet the designed quality standard, nonetheless positively influence the user experience. OD4, however, shows a slightly negative value (–2), pointing to a minor discrepancy in the opposite direction.

Information displays consistently positive values across all O-Ds (from +11 to +53). This indicates that users tend to perceive the availability and clarity of information as better than what the MC assessment records.

Safety & Security shows a mixed trend. OD2 registers a positive gap (+21), suggesting that users feel safer than what objective indicators would suggest. In contrast, OD1, OD3, and OD4 exhibit negative values (from –18 to –26), indicating that users feel less safe than what the infrastructure conditions objectively provide.

Finally, Time presents strongly positive gaps across all O-Ds (from +5 to +57). This means that users consistently perceive their travel times, waiting times, and overall journey efficiency as better than what the objective metrics indicate.

Overall, QG3 reveals a complex interplay between objective quality and subjective experience. Users systematically overestimate certain aspects, such as Customer Care, Information, and Time, while consistently underestimating Accessibility and Comfort. These divergences highlight the importance of integrating both subjective and objective perspectives in the quality management framework, as each dimension captures different facets of the mobility experience that cannot be inferred from the other alone.

Interpretation of QG4 (PUS – PUI)

QG4 compares users' expectations (PUI) with their actual satisfaction (PUS), revealing whether users feel that the service they receive meets, exceeds, or falls short of their anticipated quality standards. Negative values indicate that satisfaction is lower than expected, signalling a perceived deficit in quality, whereas positive values show that users are more satisfied than expected.

Across the four O-D pairs, Accessibility presents consistently negative QG4 values (from –13 to –50), indicating that users systematically receive less accessibility quality than they expect. The magnitude is particularly high in OD1 (–40) and OD4 (–50), suggesting that in these corridors, expectations about accessibility are unmet to a substantial degree. Even in OD3, where objective accessibility conditions were relatively stronger, the positive gap (–13) reveals that satisfaction still lags behind expectations.

Availability shows a mixed pattern. OD1 has a positive gap (+20), meaning users are more satisfied than expected-possibly because the availability of transport services exceeded their initial expectations. In contrast, the remaining O-Ds display large negative gaps (–62, –62, –75), indicating that expectations about coverage, service presence, or multimodal options are far from satisfied. The particularly low value in OD4 (–75) highlights a substantial discrepancy, suggesting that this corridor suffers from limited or fragmented availability from the user perspective.

Comfort presents moderate negative gaps across all O-Ds (from –20 to –31), showing that users consistently expect a higher level of comfort than what they ultimately experience. Although the discrepancies are less severe than in other criteria, the uniform positivity of the gap indicates a structural shortfall in meeting users' comfort expectations.

Customer Care shows the same pattern of expectation-satisfaction imbalance. Except for OD3, which presents a neutral gap (0), all other O-Ds show negative values (–20, –62, –25). The particularly low gap in OD2 (–62) reflects a pronounced mismatch between what users expect in terms of assistance or support and what they actually encounter. The fact that OD3 shows a perfect match between expectations and satisfaction likely derives from uniformly low levels of both, as previously noted.

Environmental Impact displays an uneven pattern, oscillating between negative and positive gaps. OD1 (+20) and OD3 (+50) show positive values, indicating that user satisfaction exceeds expectations. Conversely, OD2 (0) suggests a balance between expectations and satisfaction, whereas OD4 (–25) shows a shortfall where expectations are higher than perceived results. These fluctuations suggest that environmental perceptions vary considerably depending on the specific urban context of each corridor.

Information shows consistently negative values across all O-Ds (from –25 to –40), indicating that user expectations are not being fully met. The fact that gaps are moderately low and consistently negative across the board suggests that information provision is uniformly perceived as insufficient relative to user expectations, even if satisfaction is not extremely low.

Safety & Security presents large negative gaps in all O-Ds (from –18 to –75), signalling a notable expectation-satisfaction mismatch regarding perceived safety conditions. OD1 (–75) and OD4 (–62) reveal particularly large discrepancies, suggesting that users' safety expectations are significantly higher than the safety level they experience. Even in the

“best” scenario in OD2 (–18), expectations still exceed satisfaction, highlighting safety as a recurrent area of perceived under-quality.

Finally, Time exhibits a highly polarised configuration. OD1 shows a strongly positive value (+60), meaning that users are substantially more satisfied with travel time and waiting times than expected, likely because their expectations were low to begin with. OD2 shows perfect alignment between expectations and satisfaction (0), while OD3 displays a sizeable negative gap (–38), suggesting that users consider travel or waiting times less favourable than anticipated. OD4 also reflects an equilibrium (0), showing that expectations are in line with experienced conditions for time-related aspects.

Overall, QG4 reveals a general tendency for expectations to exceed satisfaction across most criteria, highlighting several criteria - particularly Accessibility, Availability (in some O-Ds), Comfort, Information, and Safety & Security - where perceived quality does not meet users’ anticipations. Only a few criteria (notably Time in some O-Ds and Environmental Impact in specific corridors) display contexts where satisfaction surpasses or matches expectations. This final component of the gap analysis provides critical insights into user priorities and areas where interventions are most urgently needed to align the mobility offer with public expectations.

*Figure 47 s*ummarises the Quality Gaps (QG) arising from the four dimensions of quality - designed, delivered, expected and perceived - along the four O-Ds analysed, and in relation to the eight main quality criteria.

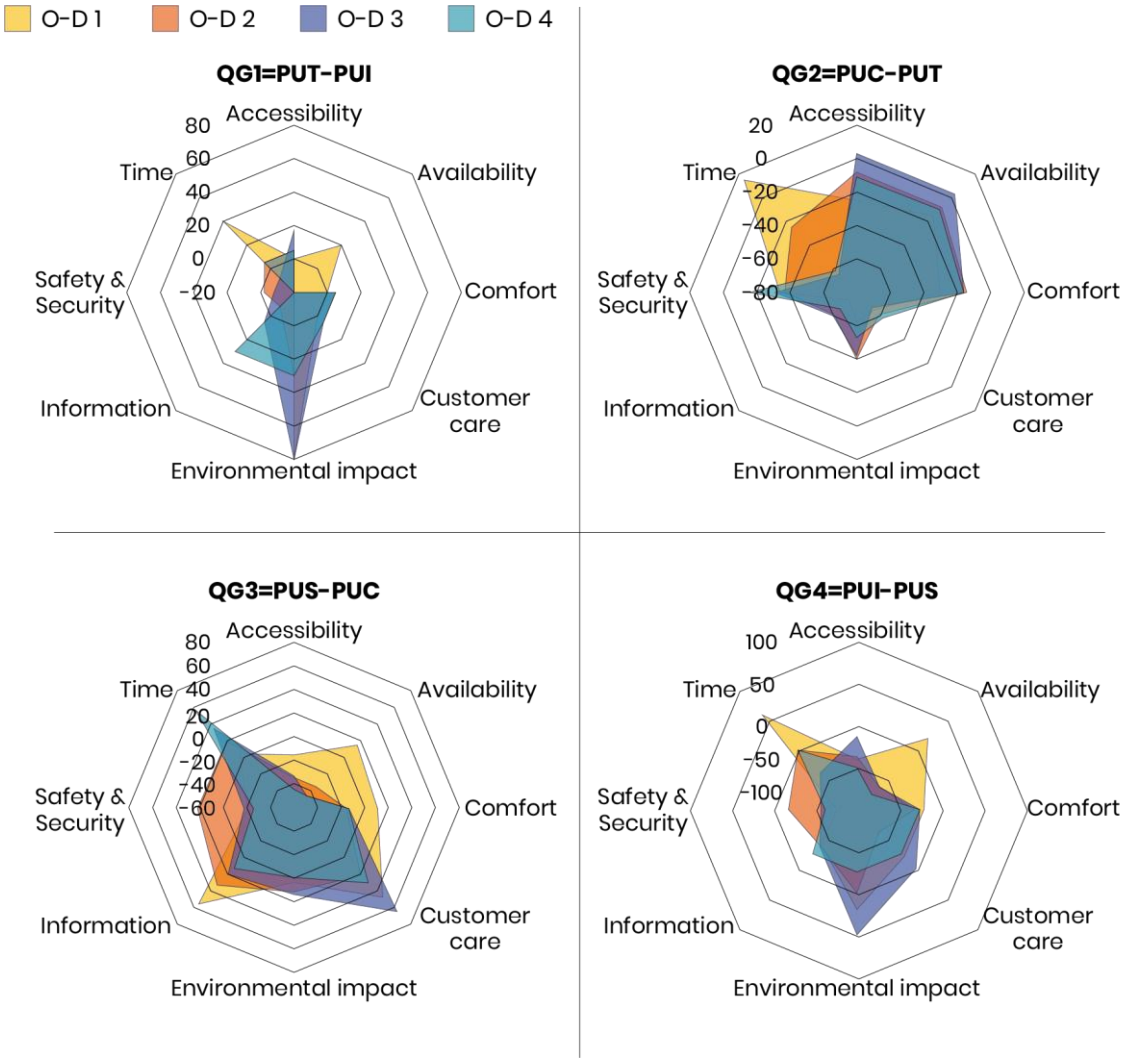


Figure 47. Quality Gaps arising from the four quality dimentions along the four O-Ds analysed.

Overall, the Quality Gap Analysis provides a comprehensive picture of how expected, perceived, designed, and delivered quality interact across the four O-D pairs. The results reveal a system characterised by structural weaknesses in the delivered quality relative to the designed standard (QG2), persistent misalignment between objective conditions and user perceptions (QG3), and in several cases a substantial divergence between what users expect and what they perceive (QG4). At the same time, the analysis highlights areas where expectations remain realistic (QG1), as well as criteria, such as Comfort or, in some cases, Availability, where user satisfaction is comparatively higher despite infrastructural shortcomings. The combined interpretation of the four gap dimensions demonstrates that challenges are multidimensional: some relate to inadequate provision (negative PUC), others to limited user awareness or mismatched expectations, and others still to systematic deficiencies in quality. This integrated diagnostic framework establishes a robust foundation for the next phase of the research, in which targeted improvement strategies are defined by linking each gap to the operational, managerial, or informational interventions required to enhance the overall multimodal mobility experience.

5.3.2 Definition of improvement strategies

The Quality Gap Analysis provides a detailed picture of how designed, delivered, expected, and perceived quality interact across the four O-D pairs. Based on these results, this section identifies targeted strategies for each corridor, differentiating between:

- Preventive measures: actions aimed at preserving or refining aspects where quality is already acceptable or above the designed threshold (i.e., small QG2, moderate QG3, or high PUC).
- Corrective measures: interventions required where significant gaps or systemic deficiencies emerge (large negative QG2, large positive QG1, strong mismatches in QG3 or QG4).

This structure ensures that investments and interventions are not uniformly distributed but rather calibrated according to the conditions observed along each O-D pair.

5.3.2.1 Preventive Measure

While several criteria require substantial improvements, the Quality Gap Analysis also highlights dimensions where quality remains relatively stable or where the gaps between designed, expected, perceived, and delivered quality are small. These areas do not necessitate major interventions but rather continuous monitoring and preservation of current quality levels.

Across the four O-Ds, preventive needs primarily concern Accessibility, Comfort, Customer Care, Environmental Impact, Information, Safety & Security, and Time, but in a more nuanced manner: the stability of these criteria does not imply uniformly high quality, but rather the absence of structural misalignment. In many cases, small or near-zero QG values reflect coherent relationships between user expectations and the designed system, or a reasonable match between perceived and delivered conditions, even if absolute quality levels remain moderate.

For example, several O-Ds show acceptable consistency in Accessibility and Comfort expectations relative to the designed standard, suggesting that users attribute importance to these criteria in a way that aligns with planning objectives. In other cases, such as Customer Care, Safety, or Environmental Impact, some micro-gaps remain small, indicating stability in specific components even where global quality is not optimal. Time quality also shows moderate stability, especially regarding perceived and delivered travel times.

Preventive measures therefore aim at preserving functional assets, preventing deterioration, and ensuring operational continuity, particularly for access points, micro-infrastructure, greenery, signage, lighting, safety elements, and travel-time reliability.

The following *Table 27* provides a consolidated overview of preventive-oriented measures, organised by criterion. These measures, as well as corrective measures, were determined based on the results of the analysis and the quality gaps that emerged.

Table 27. List of preventive measures (PM) divided by quality criteria.

Accessibility	PM01. Keep pedestrian and cycling access points unobstructed.
	PM02. Preserve compliant curb ramps and access facilities already in good condition.
	PM03. Regularly monitor access-point functionality and prevent deterioration.
	PM04. Ensure accessible paths remain uninterrupted during construction works or temporary deviations.
Availability	PM05. Maintain continuity of existing mobility services and intermodal connections.
	PM06. Ensure continued functionality of bike-sharing and shared mobility services.
Comfort	PM07. Maintain routine cleaning of pedestrian/cycling paths and LPT stops.
	PM08. Preserve existing benches, bike racks and small-scale amenities.
	PM09. Ensure routine upkeep of all micro-infrastructure elements.
Customer care	PM10. Ensure continuous operation of support channels, such as hotlines, apps and reporting systems.
	PM11. Maintain digital and physical tools used for user assistance and issue reporting.
Environmental impact	PM12. Protect existing vegetation, avoiding any loss of green coverage.
	PM13. Monitor the health and condition of existing green areas and maintain them proactively.
Information	PM14. Preserve existing signage, ensuring legibility and cleanliness.
	PM15. Maintain tactile paving and tactile maps, ensuring they remain functional.
Safety & Security	PM16. Perform routine checks on lighting infrastructure to ensure proper functioning.
	PM17. Monitor driver behaviour and traffic speeds to prevent declines in safety levels.
	PM18. Maintain existing traffic-calming elements and ensure they remain effective.
Time	PM19. Preserve current walking and cycling travel-time quality.
	PM20. Ensure stability of LPT schedules and headway regularity.

These preventive measures were developed to reinforce quality aspects that were either not identified as critical (or were found to be relatively low) in the gap analysis. Their implementation similarly to the corrective measures varies according to the specific origin-destination (O-D) pairs examined and the four quality dimensions considered: designed, delivered, expected, and perceived. In the following sections, we present a detailed analysis for each O-D pair.

O-D 1: Borgo Trento – Città Antica

For O-D 1, only a limited set of criteria shows minimal discrepancies across the different quality gaps. Although these dimensions do not fully align with the Designed Quality (PUT), they do not exhibit structural weaknesses requiring substantial corrective interventions. Instead, they call for preventive-oriented measures aimed at preserving current conditions and preventing quality deterioration over time.

Accessibility is among the most stable criteria for this corridor. The absence of divergence in QG1 indicates that users' expectations are coherent with the Designed Quality, suggesting that accessibility is perceived as appropriately prioritised. Preventive efforts

should therefore ensure that existing access points remain functional, unobstructed, and well-integrated with surrounding infrastructure.

Comfort also displays initial stability, with expectations fully aligned with the Designed Quality (QG1 = 0). Although other quality gaps reveal broader issues addressed in the corrective strategy section, this alignment implies that users attribute to comfort a level of importance consistent with planning intentions. Preventive measures should thus preserve the current provision of micro-infrastructure, such as seating, bins, and small amenities, and ensure that routine cleaning cycles remain effective.

In the case of Customer Care, the absence of discrepancy in QG1 reflects a strong match between the importance assigned by users and the level of service embedded in the design. Preventive measures should focus on safeguarding existing support tools, including hotlines, help points, or digital reporting systems, ensuring their continued visibility and reliability.

Environmental Impact presents overall deficiencies addressed elsewhere; however, the small positive value in QG3 indicates that perceived conditions are slightly better than measured ones, signalling a degree of stability in users' environmental experience. Preventive efforts should therefore concentrate on protecting existing vegetation and ensuring that no further reduction in green elements occurs along the corridor.

Finally, Time shows a micro-alignment between delivered and perceived conditions (QG3 = 5), suggesting that users recognise a degree of consistency in travel-time quality. Preventive measures should preserve current pedestrian and cycling travel times and ensure continued reliability in LPT operations, particularly through schedule monitoring and the avoidance of new obstacles or bottlenecks.

O-D 2: Borgo Milano – Città antica

Although O-D 2 presents substantial misalignments across several quality criteria, a limited set of dimensions shows comparatively stable behaviour and therefore requires preventive-oriented measures rather than full corrective restructuring. These criteria display either a partial alignment between Designed, Delivered, Expected and Perceived Quality, or micro-gaps that indicate relative stability in user experience.

Accessibility demonstrates a near-alignment between user expectations and design intentions (QG1 = 5), suggesting that the importance users attribute to accessibility corresponds reasonably well with planning priorities. Preventive measures should ensure that existing accessible elements, such as ramps, entry points, and transitions, remain functional and unobstructed.

Comfort, despite exhibiting operational issues in other gaps, shows a near-zero discrepancy between expectations and design (QG1 = –1), indicating stable user relevance. Preventive measures should focus on preserving current micro-infrastructure

(e.g., benches, resting areas, bike racks) and maintaining regular cleaning cycles to prevent deterioration.

Environmental Impact shows a small positive perception-delivery micro-gap (QG3 = +10), meaning users perceive the environmental conditions as slightly better than what the objective assessment reports. Preventive measures should thus protect existing greenery and microclimatic features along the corridor, preventing further loss of environmental quality. At the same time, the expectation-satisfaction relationship is stable (QG4 = 0), suggesting that users' expectations are fully aligned with their perceived experience. Given this overall stability, preventive measures should focus on preserving existing greenery and microclimatic features along the corridor, ensuring that the current environmental quality is maintained and that no degradation occurs over time.

Time shows a perfect alignment between expectations and satisfaction (QG4 = 0), indicating that users' experience of travel and waiting times matches exactly what they anticipate. Expectations are also consistent with the designed relevance of the criterion (QG1 = +5), suggesting that time-related aspects are perceived as important in a way that reflects planning intentions. Preventive measures should therefore focus on maintaining the current quality of travel times to preserve the quality level.

Finally, Safety & Security presents a very small expectation-design divergence (QG1 = –2), meaning that the importance users attribute to safety is almost perfectly aligned with what the system is intended to provide. This near-zero misalignment indicates stability in the perceived relevance of safety within the mobility experience. As such, preventive measures should ensure that existing lighting and traffic-calming elements remain functional, safeguarding the current balance between system intent and user expectations.

O-D 3: Valdonega - Cittadella

Despite the presence of several discrepancies across multiple quality criteria, O-D 3 also presents a number of dimensions where gap values are modest or display relative stability. These criteria do not require extensive corrective interventions but instead call for preventive-oriented measures aimed at preserving existing quality levels.

Accessibility shows moderate alignment between user expectations and perceived conditions (QG4 = –13), meaning that users' experience remains reasonably consistent with what they consider acceptable. Since delivered quality slightly exceeds design intentions (QG2 = +3), the existing infrastructure shows structural stability. Preventive efforts should therefore prioritise keeping accessible routes unobstructed and ensuring that ramps retain full functionality.

Availability also benefits from a small positive designed–delivered alignment (QG2 =+ 3), indicating that the current multimodal offer performs marginally better than originally planned. Although perception remains below delivered conditions, the underlying

network demonstrates robustness. Preventive measures should focus on preserving existing multimodal continuity and avoiding new discontinuities, especially in walking and cycling infrastructure.

Comfort presents moderate stability, with expectations closely aligned with the designed standard (QG1 = +5), and a balanced relationship between perceived and delivered comfort. This suggests that amenities, cleanliness and micro-infrastructure are performing adequately. Preventive measures should therefore concentrate on routine cleaning, upkeep of existing equipment, and preservation of resting points and shading elements.

Customer Care shows strong coherence between expectations and design (QG1 = +5) and full alignment between expectations and satisfaction (QG4 = 0). Users therefore experience a level of support that matches what they anticipate. Preventive actions should ensure continuity of current support channels, such as hotlines, and maintain current staff presence at key nodes.

Environmental Impact displays a small positive perception-delivery micro-gap (QG3 = +12), indicating that users see environmental conditions as slightly better than what the objective assessment reports. Preventive efforts should protect existing vegetation.

Information shows coherent expectations relative to design (QG1 = +5). This indicates that signage, tactile elements and wayfinding systems still provide functional support. Preventive measures should therefore focus on preserving legibility, monitoring wear and tear, and ensuring that tactile paving and other accessibility-related informational features remain operational.

Safety & Security presents a not entirely stable relationship between delivered conditions and perception (QG3 = –18), while expectations show near-alignment with design (QG1 = –8). While safety remains a critical dimension overall, this micro-alignment suggests that users' perception is consistent with real conditions. Preventive actions should include regular functional checks of lighting and routine monitoring of vehicle speeds and potential conflict points.

Time exhibits perfect expectation-satisfaction alignment (QG4 = 0), indicating that users experience travel and waiting times exactly as expected. The slight expectation-design misalignment (QG1 = –8) shows stable relevance perceptions. Preventive measures should therefore focus on maintaining current travel-time predictability for walking, cycling and LPT, ensuring that no additional sources of delay or inefficiency are introduced.

O-D 4: Borgo Venezia – Città antica

Although O-D 4 is characterised by several substantial discrepancies across multiple quality dimensions, a number of criteria display minimal or near-zero gap values. These

more stable dimensions do not require extensive corrective action and should instead be supported through preventive-focused strategies to preserve their current quality.

Time shows a stable alignment between expected and perceived quality (QG4 = 0). This indicates that users' expectations are coherent with their actual experience along the corridor. Preventive efforts should therefore aim to safeguard current travel-time conditions by ensuring continuity in pedestrian and cycling paths and monitoring LPT schedule adherence.

For Environmental Impact, although overall quality remains weak, the perception-delivery relationship (QG3 = –2) reveals partial stability. Users tend to perceive environmental conditions in a manner that roughly corresponds to the actual situation. Preventive interventions should thus emphasise the protection of existing vegetation and the preservation of micro-green elements to avoid further environmental decline.

In O–D 4, several criteria show a small positive alignment between user expectations and the designed level of service (QG1 = +5 for Accessibility, Comfort and Customer Care). This indicates that the importance users attribute to these dimensions closely matches the priorities embedded in the planning and design of the corridor. Such consistency suggests that the relevance of these criteria is well understood by service providers and appropriately reflected in the system's intended quality standards although other gaps highlight structural or perceptual shortcomings. Preventive measures should therefore focus on maintaining this alignment.

5.3.2.2 Corrective Measure

The analysis of the four selected O-D pairs reveals a recurring pattern of quality deficiencies that require targeted corrective interventions. Although the magnitude of the gaps varies across the network, several criteria consistently emerge as critical: Accessibility, Comfort, Customer Care Environmental Impact, Information, and Safety & Security. These dimensions show large and persistent discrepancies between Designed Quality (PUT), Expected Quality (PUI), Perceived Quality (PUS), and Delivered Quality (PUC), as reflected in strongly negative QG2 and QG3 values and substantial divergences in QG1 and QG4.

Environmental Impact and Information are the most structurally critical criteria across all O-Ds: nearly all corridors exhibit insufficient green infrastructure. Safety & Security also shows significant mismatches, particularly where perceived safety falls far below expectations, signalling discomfort even when infrastructure is partially compliant. Accessibility and Comfort present additional recurrent gaps, typically linked to physical discontinuities along walking and cycling routes, limited amenities, and localised barriers that reduce the usability of the multimodal network. Customer Care issues are widespread as well, particularly where delivered assistance and support services fail to meet both user expectations and the designed standard.

Corrective measures therefore focus on addressing these systemic weaknesses through structural upgrades, expanded green and informational infrastructure, enhanced lighting and traffic-calming strategies, strengthened intermodal integration, and improved user support mechanisms.

The following *Table 28* synthesises all recommended corrective actions, grouped by criterion.

Table 28. List of corrective measures (CM) divided by quality criteria.

Accessibility	CM01. Design bicycle-access ramps at all transitions between cycle lanes, sidewalks, and carriageways, ensuring continuity with existing or planned infrastructure and compliance with universal design standards.
	CM02. Standardise curb ramps at pedestrian crossings, reducing step heights and uneven surfaces, particularly for users with reduced mobility.
	CM03.Address localised accessibility gaps caused by steep gradients or irregular pavement conditions.
	CM04. Reinforce intermodal access points, ensuring seamless coordination between walking, cycling and public transport.
Availability	CM05. Complete missing segments of the cycling network, especially near intersections and fragmented sections.
	CM06. Improve path connectivity by adding safer crossings and more direct pedestrian/cycling routes.
	CM07. Expand bike-sharing services by increasing the number and density of docking stations.
	CM08. Extend shared-mobility services to underserved areas.
Comfort	CM09. Enhance cleaning operations, including more frequent litter removal and systematic graffiti removal.
	CM10. Increase seating availability, especially at LPT stops and along long pedestrian sections.
	CM11. Expand the number of public toilets, ensuring proper maintenance, signage, and accessibility.
	CM12. Add bicycle racks, prioritising sheltered or covered models for secure intermodal transitions.
	CM13. Increase the number of drinking fountains, especially in high-pedestrian-flow segments.
Customer care	CM14. Increase the presence of LPT personnel at key stops and transfer nodes.
	CM15. Strengthen assistance channels, improving visibility and usability of hotlines, apps, QR-based reporting systems and info points.
	CM16. Provide training for LPT staff and municipal personnel on overall accessibility, multimodal guidance and user interaction.
Environmental impact	CM17. Expand green infrastructure, adding trees, shrubs, linear planters, or pocket parks in heat-exposed or impervious areas.
	CM18. Reinforce ecological continuity by linking isolated green patches into continuous green corridors.
	CM19. Integrate vegetation along pedestrian and cycling routes to improve microclimatic comfort.
	CM20. Implement targeted micro-greening interventions in highly paved areas.
	CM21. Add natural or artificial shading elements along exposed walking and cycling segments or at uncovered LPT stops.
Information	CM22. Install clearer and more frequent wayfinding elements along walking, cycling and LPT routes.
	CM23. Improve accessibility-related signage, including indications of ramps, barrier-free paths, and path continuity.
	CM24. Strengthen tactile maps and tactile paving, ensuring visibility, and full functionality.
	CM25. Standardise signage style and layout to improve continuity along the corridor.
	CM26. Add panels showing accessibility, intermodal transfer points, and LPT stop conditions (e.g., ramp locations, platform heights, step-free routes).

Safety & Security	CM27. Improve street lighting, particularly in isolated pedestrian and cycling segments and at intersections.
	CM28. Implement traffic-calming measures, such as raised crossings, reduced-speed zones, chicanes and narrowed lanes.
	CM29. Implement policies that discourage private car use in favour of more sustainable modes of transport.
Time	CM30. Improve LPT reliability, strengthening headway management, schedule coordination and traffic-signal priority.
	CM31. Complete missing cycling segments to ensure direct and continuous routes.
	CM32. Reduce pedestrian delays by adding crossings and removing unnecessary barriers.
	CM33. Provide more direct walking and cycling shortcuts, especially where current routes are indirect.

These corrective measures were developed to address the critical issues highlighted by the gap analysis. Their implementation varies according to the specific O-D pairs examined and the four qualitative dimensions considered: designed, delivered, expected, and perceived quality. In the following sections, we provide a detailed analysis of the main quality gaps identified for each O-D pair.

O-D 1: Borgo Trento – Città Antica

The Quality Gap Analysis indicates that O-D 1 is characterised by substantial misalignments across several quality dimensions. The most critical discrepancies emerge in Environmental Impact, Information, Safety & Security, Customer Care, and Accessibility, with Comfort also showing notable, though comparatively less dramatic, inconsistencies. These results suggest the presence of structural deficiencies and significant divergences between what users expect, what the system is designed to provide, and what is effectively delivered in practice.

Accessibility displays one of the clearest signs of misalignment, with a substantial gap between user expectations and perceived experience (QG4 = –40), and a considerable mismatch between delivered and designed quality (QG2 = –25). This combination points to both persistent physical barriers and a generalised sense of user dissatisfaction.

For Comfort, the gap between expected and designed quality is minimal (QG1 = 0), yet delivered quality remains consistently below acceptable levels (QG2 = –30; QG3 = +10). This indicates that shortcomings arise from the actual condition of infrastructure and amenities rather than from user perception or unrealistic expectations.

Customer Care also reveals marked inconsistencies. Users report uneven or insufficient assistance, while delivered quality falls significantly below the intended standard (QG2 = –65; QG3 = +45), highlighting weaknesses in both operational support and user interaction.

A similar pattern is observed for Environmental Impact, one of the weakest-performing criteria along this O-D. The combination of very high expectations relative to design (QG1 = +80) and a large negative gap between delivered and designed quality (QG2 = –62)

signals a systemic deficit in environmental conditions, including limited vegetation, shading, and ecological continuity.

Among all criteria, Information stands out as particularly problematic. Delivered information is far below the designed standard (QG2 = –73), while perceived information is well below user expectations, as suggested by the positive QG3 and QG4 results. This indicates that users struggle to access clear, consistent, and accessible information throughout the corridor.

Safety & Security also exhibits severe misalignments. The exceptionally large expectation-perception gap (QG4 = –75) highlights users' strong sense of discomfort, even in cases where objective safety conditions may not be critically compromised. This suggests that both actual and perceived safety require substantial improvement.

Although not as critical as the dimensions described above, Availability and Time still present non-negligible gaps that justify dedicated corrective actions. Availability suffers from repeated negative values in QG2 and QG4, reflecting persistent difficulties in aligning delivered services with user expectations. Time generally performs better, yet specific gaps, particularly the perception-expectation misalignment captured by QG4, indicate that travel-time reliability could benefit from targeted interventions to prevent future deterioration.

O-D 2: Borgo Milano – Città Antica

The Quality Gap Analysis for O-D 2 highlights significant misalignments across multiple criteria, although the intensity and distribution of these gaps differ from those observed along O-D 1. The most critical shortcomings emerge in Information, Customer Care, Environmental Impact, and Safety & Security, while Accessibility, Availability, and Comfort also display relevant, albeit less severe, inconsistencies. Overall, the results point to both infrastructural weaknesses and considerable divergences between what users expect, what is planned, and what is actually provided.

Accessibility presents moderate but meaningful discrepancies across all four gap dimensions (QG1 = +5; QG2 = –8; QG3 = –34; QG4 = –37). These values indicate that users expect a higher level of accessibility than what is currently delivered, and that perceived conditions fall short of both expectations and the designed standard.

A similarly complex pattern emerges in Availability, where the system exhibits notable misalignments between designed and delivered quality (QG2 = –8) and a clear mismatch between perceived conditions and user expectations (QG4 = –62). The high expectation-perception gap suggests that users experience a scarcity of available services or infrastructure, despite their greater importance in this corridor.

Comfort also shows moderate inconsistencies (QG2 = –14; QG3 = –16; QG4 = –31), reflecting a situation in which users perceive comfort levels to be consistently lower than

both the designed and delivered conditions, thus signalling dissatisfaction with micro-infrastructure, amenities, or environmental characteristics.

Among the most problematic dimensions, Customer Care presents pronounced disparities, revealing persistent difficulties in the visibility, responsiveness, and perceived effectiveness of support services. These gaps indicate that users often feel insufficiently assisted during their journeys, even when operational standards may appear satisfactory on paper.

Environmental Impact reveals a similarly challenging condition, with marked discrepancies between high user expectations (QG1 = +30) and low delivered quality (QG2 = –40), while perceived environmental conditions remain modestly above delivered levels (QG3 = +10). This suggests a corridor characterised by limited vegetation, exposure to heat, or insufficient green infrastructure, contributing to an overall sense of environmental degradation.

The most severe misalignments, however, concern Information, which stands out as one of the weakest-performing criteria along O-D 2. All gap dimensions point to deep inconsistencies: delivered information is far below both the designed level (QG2 = –61) and user expectations (QG3 = +31; QG4 = –38). This indicates that users lack a clear, accessible, and reliable wayfinding and service-related information, affecting both navigation and multimodal integration.

Finally, Safety & Security also displays significant discrepancies (QG2 = –37; QG3 = 21; QG4 = –18), suggesting that perceived safety remains considerably lower than the objective conditions, while delivered safety measures fail to match design standards.

Time, although less problematic, still shows moderate misalignments (QG2 = –25; QG3 = 20), indicating that perceived travel-time reliability does not fully align with the quality measured on the ground.

O-D 3: Valdonega – Cittadella

The Quality Gap Analysis for O-D 3 reveals a heterogeneous but consistently problematic pattern of misalignments across several criteria. Although this corridor performs marginally better than others in certain structural aspects, users still experience substantial dissatisfaction across key dimensions, particularly Environmental Impact, Safety & Security, Information, Accessibility, Availability, and Comfort. These results indicate that both infrastructural limitations and perceptual misalignments contribute to an overall decline in quality.

Accessibility exhibits clear discrepancies between what users expect, what is designed, and what is ultimately perceived. Expectations exceed the designed standard (QG1 = +17), delivered quality is only marginally above the design benchmark (QG2 = +3), and yet perceived conditions remain far below actual quality (QG3 = –33), producing a significant

expectation-perception gap (QG4 = –13). This suggests that structural accessibility is uneven along the corridor and that many improvements remain largely unnoticed or ineffective from the user's point of view.

Availability presents an even more polarised pattern. While delivered availability slightly exceeds what was planned (QG2 = +3), both QG3 (–45) and QG4 (+45) reveal severe discrepancies between delivered conditions and user perceptions. This indicates that the existing service offer fails to be interpreted as adequate or accessible, highlighting a serious perceptual deficit.

Customer Care shows some of the largest misalignments recorded among all corridors. Delivered quality is dramatically below the designed level (QG2 = –67), and users judge the actual assistance received far more negatively than what the objective assessment suggests (QG3 = +62). These discrepancies point to both operational inconsistencies and substantial weaknesses in perceived responsiveness and support.

Environmental Impact emerges as one of the weakest-performing criteria. Users expect significantly better environmental quality (QG1 = +80), yet delivered quality is far below this standard (QG2 = –42). Even though perceived environmental conditions slightly exceed delivered ones (QG3 = +12), the overall picture reflects a corridor lacking sufficient vegetation, shading, or ecological continuity.

Similarly, Information shows deep structural deficiencies. Expectations are slightly above the designed standard (QG1 = +5), but delivered information falls dramatically short (QG2 = –60). The positive perception-delivery gap (QG3 = +18) shows that users perceive information as marginally better than what is measured, yet still inadequate compared to expectations (QG4 = -37). This confirms widespread issues in wayfinding, accessibility signage, and multimodal guidance.

Safety & Security displays significant misalignments as well. While delivered conditions are not critically poor, they do not reach the designed level (QG2 = –31), and perceived safety is substantially lower than both (QG4 = -57). This suggests persistent feelings of vulnerability along the corridor, possibly exacerbated by limited lighting, and traffic exposure.

Finally, Time shows notable misalignments between design and delivery (QG2 = –65), although users perceive travel times as somewhat better than delivered conditions suggest (QG3 = +35). This reflects inconsistencies in infrastructure continuity and delays generated at intersections or LPT stops.

While O-D 3 exhibits fewer extreme conditions than other corridors, several criteria (including Accessibility, Availability, Comfort, Environmental Impact, and Information) still present non-critical but persistent gaps that require proactive attention to avoid further deterioration.

O-D 4: Borgo Venezia – Città antica

The Quality Gap Analysis for O-D 4 highlights several pronounced misalignments across multiple quality dimensions, revealing a corridor in which user expectations, design intentions, and delivered quality diverge considerably. Severe discrepancies appear in Availability, Accessibility, Information, Safety & Security, Customer Care, and, to a lesser extent, Comfort and Environmental Impact. Although some dimensions show moderate stability, the overall quality suggests the need for substantial corrective action.

Accessibility presents significant inconsistencies in all gap dimensions (QG1 = +5; QG2 = –11; QG3 = –44; QG4 = –50). These values indicate that delivered quality falls short of the design standard and that users perceive accessibility as dramatically lower than both designed and delivered conditions. The large expectation-perception gap reflects persistent structural barriers and insufficiently visible improvements.

Availability is one of the most critical dimensions in this corridor. All four gap indicators show substantial misalignments (QG1 = –20; QG2 = –10; QG3 = –45; QG4 = –75), suggesting that users expect a far denser and more continuous service provision. Delivered availability remains below design intentions, and perceived conditions fall even further behind.

Comfort also displays moderate but meaningful discrepancies (QG2 = –16; QG3 = –14; QG4 = –25). This indicates inadequacies in micro-infrastructure, seating, shading, or cleanliness, which cumulatively reduce perceived comfort along the corridor.

Customer Care shows notable inconsistencies (QG2 = –58; QG3 = 28; QG4 = –25), reflecting weak alignment between designed and delivered service quality and user dissatisfaction with perceived support and information availability.

Environmental Impact presents persistent but less severe misalignments (QG1 = 30; QG2 = –53; QG4 = –25). Users expect greener environments, but delivered conditions fall noticeably below these standards, with limited vegetation and insufficient ecological integration.

Information emerges as one of the weakest-performing criteria for O-D 4. Delivered quality is significantly below the designed standard (QG2 = –66), and users perceive information as inadequate (QG3 = 11; QG4 = –25). This indicates major deficiencies in wayfinding, accessibility indications, multimodal connectivity information, and the visibility of tactile or visually accessible elements.

Safety & Security is also a critical concern in this corridor. Delivered safety falls below design intentions (QG2 = –16), users perceive safety levels as poor (QG3 = –26), and the difference between expectations and perceptions is extremely large (QG4 = –62). This suggests discomfort due to traffic exposure, poor lighting, and conflict points.

Time, while less problematic, still shows moderate misalignments (QG2 = –62; QG3 = 57). Delivered travel times are significantly worse than what was designed, although users perceive them as somewhat more acceptable, likely due to familiarity with the corridor.

Beyond the priority criteria, several dimensions (including Comfort, Environmental Impact, Availability, and Information) present sub-critical but persistent gaps that require targeted interventions to prevent the progressive decline of overall service quality.

5.3.3 Stakeholder roles in the implementation of quality measures

The effective implementation of the identified quality measures requires the coordinated involvement of multiple institutional and operational actors, each with specific skills and responsibilities (*Figure 48*). Preventive and corrective measures, although differing in scope, intensity, and time horizon, share the need for a multi-level governance framework capable of ensuring continuity, coherence, and integration across the different quality criteria considered.

The **Public Administration** plays a central role, acting as the primary strategic and coordinating body. Through its relevant departments (mobility, public works, urban planning, environment, and so on), the municipality is responsible for planning activities, defining quality standards, and integrating the proposed measures into broader urban policy and regulatory frameworks. In particular, the municipal administration oversees measures related to accessibility and availability of infrastructure, such as the standardisation of curb ramps, the completion of cycling networks, and the continuity of pedestrian routes, as well as the adoption of demand-management and traffic-regulation policies that support safety, environmental impact, and travel-time quality objectives.

Municipal technical offices and road infrastructure managers are the primary operational actors responsible for the design, implementation, and maintenance of physical elements within the public realm. Their tasks include the delivery of measures affecting the material quality of space, such as pavement upkeep, installation and maintenance of street furniture, lighting systems, signage, traffic-calming devices, and micro-greening interventions, as well as the protection of existing vegetation. These actors are crucial both for preventive actions—through routine monitoring and maintenance, and for corrective interventions aimed at addressing identified deficiencies, discontinuities, or deterioration.

Urban green and environmental services contribute to the implementation of measures related to environmental impact and comfort. These actors are responsible for protecting and maintaining existing green assets, expanding green infrastructure, reinforcing ecological continuity, and implementing shading and microclimatic mitigation solutions along pedestrian and cycling corridors and at exposed public transport stops.

Local Public transport operators hold a key responsibility with respect to availability, time, and customer care criteria. They are tasked with ensuring service continuity and reliability,

maintaining stable timetables and headways, and guaranteeing the proper functioning of stops and transfer nodes. In addition, they play a fundamental role in user assistance through the presence of staff at key locations, the management of support and reporting channels (e.g., hotlines, apps, and digital platforms), and the training of both operational and front-line personnel on accessibility, multimodal guidance, and user interaction.

Alongside public transport operators, **shared-mobility providers** (including bike-sharing and other shared-mobility systems) are responsible for the quality and continuity of complementary mobility services. Measures aimed at maintaining, expanding, or spatially rebalancing these services require close coordination with public authorities, both in preventive terms, ensuring the full functionality of existing fleets and docking stations; and in corrective terms, such as extending coverage to underserved areas or increasing station density in high-demand locations.

Finally, a transversal role is played by **users and civil society actors**, including representative associations (e.g., organisations advocating for persons with disabilities or active mobility). While not directly responsible for implementation, these stakeholders contribute through the reporting of critical issues, participation in monitoring processes, and the assessment of perceived quality. Customer care tools and reporting systems, therefore, function as key interfaces between institutional actors and end users, supporting continuous feedback and adaptive management.

Taken together, these stakeholders form a system of shared responsibilities in which the quality of urban multimodal transport does not depend on isolated actions, but on the capacity to coordinate preventive and corrective measures throughout the life cycle of infrastructures and services, in alignment with the quality criteria adopted as the theoretical and operational framework.

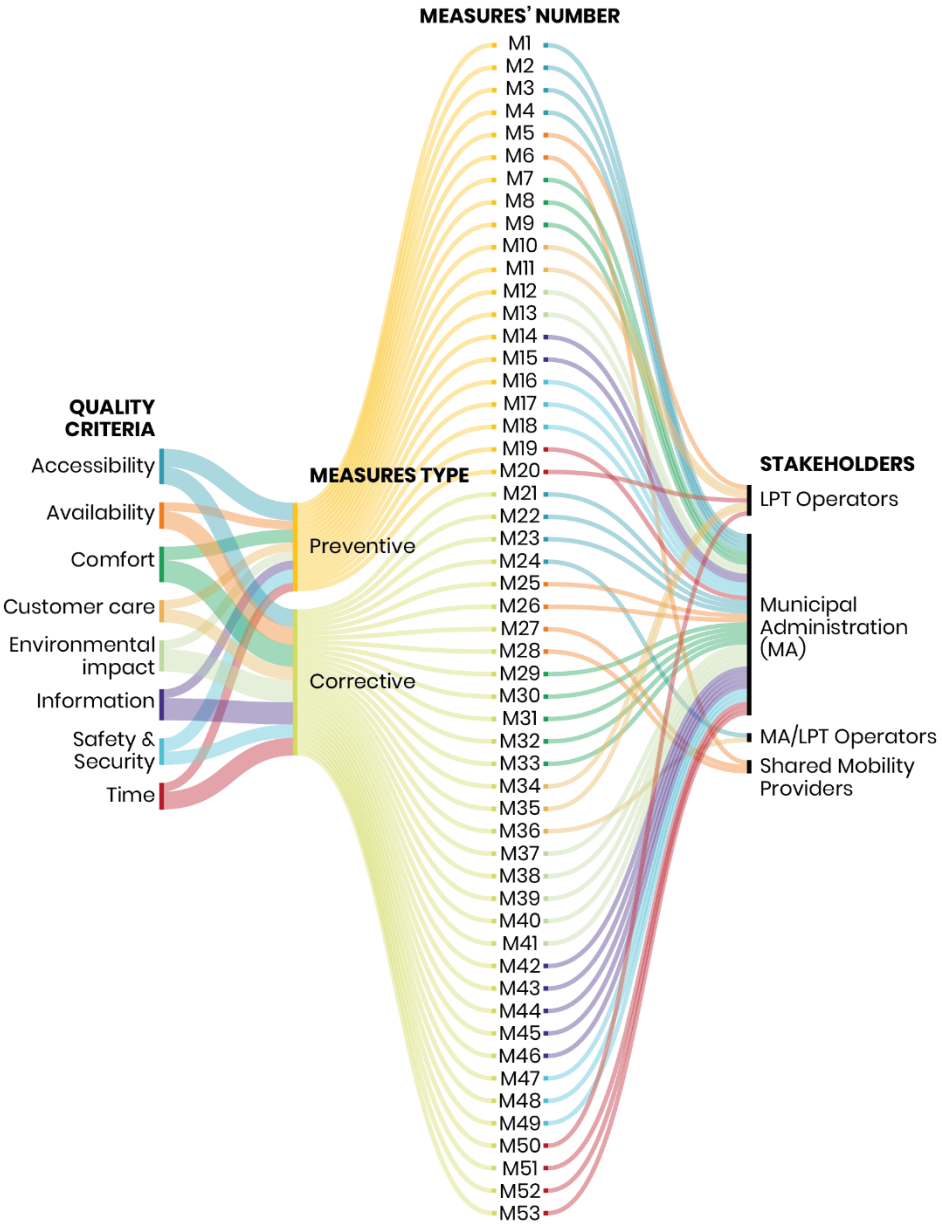


Figure 48. Alluvial diagram illustrating the relationships between quality criteria, type of measure (preventive or corrective), individual measures and responsible stakeholders in the urban multimodal transport system.

5.3.4. Concluding Remarks

The Quality Gap Analysis across the four selected O-D pairs reveals a complex and heterogeneous quality landscape, where systemic weaknesses coexist. Although the magnitude and nature of the gaps vary across corridors, several common patterns emerge. Information, Environmental Impact, Safety & Security, Availability, and Accessibility consistently exhibit the greatest misalignments between Designed, Delivered, Expected and Perceived Quality. These criteria therefore represent the most urgent priorities for intervention, as they reflect both structural deficiencies and unmet user expectations.

At the same time, other criteria such as Time, Comfort, and Customer Care show moderate discrepancies or partial stability, depending on the O-D pair. These areas require mainly preventive-oriented measures, aimed at preserving existing quality levels and preventing further deterioration. The coexistence of preventive and corrective needs illustrates the multidimensional nature of urban mobility quality, where different components evolve at different speeds and are perceived differently by users.

The improvement strategies formulated for each O-D pair translate the gap results into targeted, context-sensitive measures, distinguishing between preventive and corrective interventions. The first address the most critical deficiencies, while the second consolidate acceptable or stable conditions. This approach ensures that strategic recommendations remain proportionate, evidence-based, and aligned with users' mobility experiences.

Overall, the analysis demonstrates the value of integrating Delivered, Expected, Perceived, and Designed Quality into a unified evaluative framework. By doing so, it becomes possible not only to identify where and why quality gaps emerge, but also to prioritise interventions in a manner that is both analytically robust and operationally meaningful. These insights provide a solid foundation for future planning and design decisions, contributing to more inclusive, efficient, and user-centred multimodal mobility systems.

6. FINAL DISCUSSION & CONCLUSIONS

This PhD thesis responds to the current absence of a coherent and widely applicable framework for evaluating the quality of multimodal urban transport systems. The systematic review of the literature revealed that, although transport quality has been examined across many studies (Eboli & Mazzulla, 2010), research efforts remain compartmentalised. Most analyses concentrate on a single transport mode, typically public transport, while walking, cycling, shared micromobility, and the multimodal travel chain as a whole are often treated in isolation or overlooked entirely. As a consequence, existing assessment practices struggle to account for the interactions between modes or to describe the overall experience of urban mobility from the origin to the final destination (Moodley & Venter, 2022).

The concept of quality itself emerges as broad and multi-layered, influenced by the choice of transport mode under study, the type and origin of available data, the balance between subjective and objective perspectives, and the involvement of different institutional actors and users (Ojo, 2017). While the first quality assessment tools were developed for public transport, particularly within the Quality of Service tradition (Dell'Olio et al., 2017), contemporary research has gradually widened its scope to include walking, cycling, and, more recently, micromobility (Ferretto et al., 2024). Nevertheless, even the more advanced frameworks, such as LOS, SERVQUAL, or modal LOS variants like PLOS and MMLOS, continue to adopt modal perspectives and do not provide a fully integrated representation of multimodal quality.

The literature review highlighted two fundamental conceptual gaps that hinder the development of such an integrated framework. First, there is no commonly accepted definition of multimodal transport quality that aligns with sustainability objectives, reflects the roles and expectations of all relevant actors, and captures the wide range of attributes shaping the user experience. A comprehensive definition must encompass technical, social, environmental, technological, and experiential components, and explicitly consider the four complementary dimensions of quality: expected, designed, delivered, and perceived. Second, the absence of a shared methodological structure limits the ability to evaluate multimodal systems consistently and to compare results across contexts.

From a methodological standpoint, the field is transitioning from single-mode and technically focused evaluations to mixed approaches making use of MCDA, structural modelling, GIS-based tools, and increasingly data-driven methods. This diversification of methods reflects the growing complexity of transport systems, but also contributes to a proliferation of criteria, parameters, and indicators that vary widely across studies. The resulting heterogeneity complicates the identification of a coherent set of essential parameters for multimodal settings. At present, no universally recognised procedure exists to select, filter, or prioritise quality parameters, leading to practices that are either

overly exhaustive and thus difficult to implement, or overly narrow, failing to represent the complete travel chain.

A further gap relates to the limited use of percentage-based metrics capable of linking user expectations, design targets, delivered performance, and perceived outcomes (Barabino & Di Francesco, 2016). Such metrics are crucial for identifying quality gaps in a user-centred way and for supporting strategic decision-making in multimodal environments.

Despite these limitations, the literature shows a clear convergence toward certain foundational elements. Across different modes and methodological approaches, quality is consistently described as multidimensional, involving both physical and experiential components. Moreover, eight quality criteria emerge as recurrent and widely recognised, building on European standards (EN 13816:2002; EN 15140:2006): accessibility, availability, comfort, customer care, environmental impact, information, safety & security, and time. These criteria guarantee the representativeness of different thematic clusters linked to the quality of transport and represent a conceptual backbone for the development of a shared and robust framework for multimodal quality assessment.

By examining how the concept of quality has evolved, the thesis identifies the need for approaches that connect technical performance with user-centred perspectives. It addresses this gap by proposing a unified conceptualisation of multimodal quality, based on the multidimensional nature of quality highlighted in the literature and structured using the Logical Framework Approach (LFA).

Building on this foundation, the thesis develops a new Multimodal Quality Assessment Framework for urban environments. The framework integrates a revised definition of multimodal quality and includes the characterisation, measurement, and management of transport quality. It relies on objective measurements collected through the Mystery Client technique, subjective measurements obtained through Customer Satisfaction technique, a structured Quality Gap Analysis, and a systematic process for defining preventive and corrective measures.

The framework represents the main outcome of the research, offering a rigorous and transferable methodology for assessing and improving multimodal urban mobility.

Finally, the case study of Verona provides the practical context in which the proposed framework was tested and validated.

6.1 Advancing Quality Assessment Through an Integrated Framework

A first key contribution of this thesis is the new definition of multimodal quality, based on all the crucial points that have emerged from the literature review regarding transport quality in a multimodal perspective.

Multimodal quality is defined as the ability of an integrated multimodal transport system to meet the needs of current and potential users throughout the entire travel experience. This definition goes beyond operational performance and incorporates technical, social, environmental, and technological dimensions. It adopts a holistic perspective that integrates infrastructure and services, the roles of transport providers and public authorities, and the expectations and behaviours of users, combining subjective perceptions with objectively measurable parameters. This holistic view reinforces the conceptual link between transport planning and spatial planning.

This definition lays the foundation for future studies by defining the crucial points that should be included in multimodal quality assessment to be comprehensive, scalable, and applicable across diverse contexts.

Likewise, the proposed Multimodal Quality Assessment Framework defines the fundamental steps to follow to evaluate multimodal quality in a standardized way, integrating multiple sustainable transport systems, including LPT, non-motorized mobility (pedestrian and bicycle), and electric micromobility; combining objective measurements with user experiences; including aspects of quality that reflect multiple thematic clusters beyond the operational, geometric, and physical ones, such as social, psychological, environmental, and aesthetic ones.

Focusing on the first phase of the quality characterization, key contribution of this thesis is the development of an algorithmic methodology for deriving a set of KQPs. Existing studies typically select parameters ad hoc or based on specific modalities, leading to inconsistent comparability across modalities and contexts. In this case, the selection process combines expert judgement, AHP, Monte Carlo simulations to correct coverage errors, and triangulation with bibliometric and frequency-based classifications. The final set of quality parameters emphasises the intrinsic link between urban design and transport planning, highlighting the role of urban elements in enabling or constraining mobility. Literature consistently shows that urbanistic features such as well-designed sidewalks, barrier-free bus stops, and ramps play a crucial role in shaping accessibility and comfort (Rosa and Lopes, 2019; Podciborski, 2017; Douglas, 2015). The internal accessibility of transport modes, reflected in parameters like entrances and exits, is not only a matter of compliance with universal design principles (Chauhan et al., 2021) but also a determinant of seamless multimodal integration, as evidenced by the importance of dedicated access points for public transport and bike-sharing systems (Burlando et al., 2021; Shin, 2021). Availability further extends to spatial continuity and coverage, where the regularity of pedestrian and cycling infrastructure critically affects user experience and mode choice (Rossetti and Tiboni, 2020; Liang et al., 2021). Comfort is closely tied to the maintenance and cleanliness of these urban infrastructures, a factor linked to overall satisfaction (Barabino, 2018; Arellana et al., 2020; Sousa et al., 2019). Amenities such as benches, public restrooms, and drinking fountains contribute to comfort by addressing basic human needs during travel, thus enhancing the attractiveness of active and shared mobility options (Soltani et al., 2017; Groenendijk et al., 2018; Ujjwal and Bandyopadhyaya,

2021). Information, particularly related to accessibility, emerges as a key enabler for all users, but especially for those with mobility restrictions, emphasising the role of clear signage and inclusive information design in facilitating independent travel (de Aquino Traldi et al., 2022; Diana et al., 2016). Safety and security are critical factors influencing both the actual risk of accidents or crime and users' perception of travel quality (Tan et al., 2007; Fistola, 2020). These dimensions affect mobility choices and comfort, underscoring the importance of designing urban environments and transport services that minimise hazards. Closely related to these aspects is the management of traffic flow, which involves balancing the interactions between pedestrians, cyclists, and motor vehicles to reduce conflicts and enhance overall safety (Rodriguez-Valencia et al., 2020; Ma et al., 2019; D'Orso and Migliore, 2020). Finally, lighting plays a fundamental role in safety, particularly in public transport stops and pedestrian paths during evening or night-time hours, contributing to a sense of security and usability in urban spaces (Jahan et al., 2020; Soltani et al., 2017; Rowangould and Corning-Padilla, 2019).

Overall, these parameters illustrate how quality in multimodal urban mobility cannot be fully understood without considering the spatial and experiential aspects of the urban environment. The interplay between infrastructure design, service provision, and user needs requires holistic planning that integrates transport with urban space to foster high-quality, sustainable mobility.

Moreover, the final set of KQPs offers a practical tool for assessing both the objective and perceived quality of multimodal urban transport systems. These parameters can support public authorities and transport providers in identifying critical issues across the full travel chain, from origin to destination, encompassing infrastructure, service provision, and user experience. Unlike traditional evaluations focused on individual modes or segments, this approach enables a holistic understanding of mobility quality, guiding targeted improvements and more integrated, user-centred planning.

The second phase of the Multimodal Quality Assessment Framework operationalises the selected KQPs by applying them to the empirical assessment of multimodal quality. In this phase, KQPs are measured through both objective (Mystery Client) and subjective (Customer Satisfaction) techniques, allowing each quality dimension to be quantified in a comparable and structured manner. This makes it possible to evaluate the actual quality of the system, to understand how users experience it, and to benchmark these results against the quality that the system was originally designed to provide. In doing so, the framework captures the relationships between designed, delivered, expected, and perceived quality, enabling a comprehensive analysis of multimodal transport quality.

Finally, the third phase of the Multimodal Quality Assessment Framework focuses on quality management and represents the main added value of the proposed methodology. By integrating the results of the four dimensions of quality (designed, delivered, expected, and perceived), the framework enables a structured analysis of quality gaps that identifies where and to what extent misalignments occur along the multimodal travel chain. Unlike

traditional evaluation tools, which often limit themselves to performance measurement, this phase translates diagnostic findings into actionable management strategies. Through a systematic process, each gap is linked to a series of corrective or preventive measures, prioritized based on their relevance, feasibility, and potential impact on the user experience. This approach allows decision makers, operators, and planners to shift from an evaluation-oriented to a management-oriented perspective, ensuring that quality is not only monitored but also actively improved. This phase thus provides a practical and transferable mechanism for integrating user needs into planning, improving multimodal system performance, and supporting policy objectives related to sustainability, accessibility, and service integration.

6.2 Recommendations for practitioners and policy-makers

The PhD thesis provides several actionable recommendations:

- adopt multimodal quality frameworks that integrate objective and subjective data, ensuring that infrastructure and service decisions reflect real user needs;
- prioritise KQPs over exhaustive indicator lists, keeping quality assessment efficient, scalable, and feasible for routine monitoring;
- institutionalise Quality Gap Analysis to identify critical gaps and guide the implementation of preventive or corrective measures over time;
- Foster multimodal governance structures capable of coordinating LPT, shared mobility, micromobility, pedestrian, and cycling infrastructure within a unified strategy.

6.3 Limitations and directions for future research

While the Quality Assessment Framework demonstrates strong potential, several limitations offer opportunities for further investigation.

First, the selection of KQPs, though statistically robust, was informed by a small sample of experts predominantly affiliated with Italian institutions, which may lead to context-specific biases. Expanding the sample to include practitioners, transport operators, shared mobility providers, and international experts would enhance the framework's generalisability.

Second, the framework focuses on urban corridors and excludes suburban or regional systems (e.g., rail-oriented travel, park-and-ride access), as well as motorised individual transport. Future applications could test the framework in diverse urban and peri-urban contexts, including cities with different mobility cultures, topographies, or governance structures.

Third, while the current framework integrates objective and subjective data, the potential of real-time information through ICT/ITS remains underexploited. Integrating continuous streams of mobility data could significantly enhance the dynamic responsiveness of quality assessment processes.

Fourth, the proposed methodology could evolve into a predictive quality model, combining the Quality Gap Analysis with machine-learning tools to anticipate future quality deficits and evaluate the impacts of planned interventions which is an important direction for Smart Mobility and data-driven planning.

Finally, one limitation concerns the case study of Verona, that is based on a restricted number of O–D pairs due to the need to conduct on-site measurements individually. This constraint affected the spatial granularity of the analysis and suggests that future applications supported by larger teams or automated data collection could expand the framework to cover a wider and more representative set of travel corridors.

6.4 Final Conclusions

This thesis advances the field of sustainable urban mobility by proposing, defining, and operationalizing a comprehensive quality assessment framework specifically designed for multimodal transportation systems. Going beyond traditional mode-specific approaches, the work addresses the conceptual, methodological, and managerial challenges that currently limit the assessment and improvement of multimodal mobility.

Conceptually, the thesis contributes to an updated and holistic definition of multimodal transportation quality, recognizing its inherently multidimensional nature. The proposed definition integrates infrastructural and operational characteristics, user expectations and perceptions, environmental conditions, and the broader experiential and behavioural dimensions that characterize the entire multimodal journey. By clarifying the relationship between designed, delivered, expected, and perceived quality, the thesis establishes a shared conceptual foundation that has long been lacking in the literature.

Methodologically, the research develops an integrated, evidence-based assessment process. This includes a rigorous method for selecting Key Quality Parameters (KQPs), based on expert judgment, analytical validation techniques, and bibliometric analysis; the combined use of objective measures collected through the Mystery Client technique and subjective measures derived from customer satisfaction surveys; and the implementation of a structured Quality Gap Analysis capable of highlighting discrepancies across all quality dimensions. The resulting framework is transferable, scalable, and suitable for both research and planning applications.

From a managerial perspective, the thesis introduces a clear and actionable process for translating analytical results into preventive and corrective measures. This management phase provides public administrations and mobility service providers with a transparent

and replicable tool for diagnosing weaknesses, prioritizing interventions, allocating resources, and monitoring improvements over time. In this way, the framework supports decision-making at both the strategic and operational levels and helps bridge the persistent gap between technical planning and user-cantered mobility.

The empirical application to the city of Verona demonstrates the framework's practical relevance and feasibility. Despite the limitations associated with analysing a small number of origin-destination pairs, the case study effectively reveals substantial quality gaps, particularly in the provision of information, safety, accessibility, comfort, and environmental conditions. Furthermore, it illustrates how aligning user expectations, perceived conditions, and delivered services can lead to more targeted and impactful mobility strategies.

Overall, this thesis makes a solid contribution to the transition towards more sustainable, inclusive, and integrated urban mobility systems. Offering a coherent conceptual foundation, a rigorous methodological framework, and a practical management tool, the research positions multimodal quality as a driver for sustainable mobility. The proposed framework supports smarter, more user-oriented planning, encouraging mobility systems that are not only technically efficient but also meaningfully aligned with the needs, behaviours, and expectations of the people they serve.

7. REFERENCES

AFNOR Certification. (2005). *NF 281—Services de transport urbain de voyageurs. Règlement de certification générique*. Saint-Denis la Plaine Cedex, France: Author.

Amoroso, S., Castelluccio, F., & Maritano, L. (2012). Indicators for sustainable pedestrian mobility. *WIT Transactions on the Built Environment*. https://doi.org/10.2495/ut120161

Arellana, J., Saltarín, M., Larrañaga, A. M., Alvarez, V., & Henao, C. A. (2019). Urban walkability considering pedestrians' perceptions of the built environment: A 10-year review and a case study in a medium-sized city in Latin America. *Transport Reviews, 40*(2), 183-203. https://doi.org/10.1080/01441647.2019.1703842

Argin, G., & Ozbil Torun, A. (2015). Walking to school: The effects of street network configuration and urban design qualities on route selection behaviour of elementary school students. In *Proceedings of the 10th International Space Syntax Symposium* (pp. 118:1-118:12). Space Syntax Laboratory, The Bartlett School of Architecture, University College London.

Ashqar, H. I., Elhenawy, M., Rakha, H. A., & House, L. (2022). Quality of service measure for bike sharing systems. *IEEE Transactions on Intelligent Transportation Systems, 23*(9), 15841-15849. https://doi.org/10.1109/tits.2022.3145669

Australian Bureau of Statistics - ABS. (2016). Population density and public transport map. https://www.arcgis.com/apps/View/index.html?appid=f937c9d1258d42f7992f748d707a88bc

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies, 1*(1), 377-386. https://doi.org/10.1162/qss_a_00019

Baek, K., Lee, H., Chung, J., & Kim, J. (2020). Electric scooter sharing: How do people value it as a last-mile transportation mode? *Transportation Research Part D: Transport and Environment, 90*, 102642. https://doi.org/10.1016/j.trd.2020.102642

Bakti, I. G. M. Y., & Sumaedi, S. (2015). P-TRANSQUAL: A service quality model of public land transport services. *International Journal of Quality & Reliability Management, 32*(6), 534-558. https://doi.org/10.1108/ijqrm-06-2013-0094

Banister, D. (2008). The sustainable mobility paradigm. *Transport Policy, 15*(2), 3-80. https://doi.org/10.1016/j.tranpol.2007.10.005

Barabino, B., & Di Francesco, M. (2016). Characterizing, measuring, and managing transit service quality. *Journal of Advanced Transportation, 50*(5), 818-840. https://doi.org/10.1002/atr.1377

Barabino, B. (2018). Automatic recognition of “low-quality” vehicles and bus stops in bus services. *Public Transport, 10*(2), 257-289. https://doi.org/10.1007/s12469-018-0180-8

Barabino, B., Cabras, N. A., Conversano, C., & Olivo, A. (2020). An integrated approach to select key quality indicators in transit services. *Social Indicators Research, 149*(3), 1045-1080. https://doi.org/10.1007/s11205-020-02284-0

Bellizzi, M. G., Eboli, L., & Forciniti, C. (2019). Segregation vs interaction in the walkways: An analysis of pedestrians’ perceptions. *Research in Transportation Business & Management, 33*, 100410. https://doi.org/10.1016/j.rtbm.2019.100410

Bellizzi, M. G., Forciniti, C., & Mazzulla, G. (2021). A stated preference survey for evaluating young pedestrians’ preferences on walkways. *Sustainability, 13*(22), 12434. https://doi.org/10.3390/su132212434

Beura, S. K., Chellapilla, H., Jena, S., & Bhuyan, P. K. (2016). Service quality assessment of shared use road segments: A pedestrian perspective. In *Advances in Intelligent Systems and Computing* (pp. 653-669). https://doi.org/10.1007/978-981-10-1645-5_55

Blečić, I., Canu, D., Cecchini, A., Congiu, T., & Fancello, G. (2016). Factors of perceived walkability: A pilot empirical study. In *Lecture Notes in Computer Science* (pp. 125-137). https://doi.org/10.1007/978-3-319-42089-9_9

Botzow, H. (1974). Level-of-service concept for evaluating public transport. *Transportation Research Record, 519*, 73-84.

Burlando, C., Ivaldi, E., & Ciacci, A. (2021). Seniors’ mobility and perceptions in different urban neighbourhoods: A non-aggregative approach. *Sustainability, 13*(12), 6647. https://doi.org/10.3390/su13126647

Calvey, J., Shackleton, J., Taylor, M., & Llewellyn, R. (2015). Engineering condition assessment of cycling infrastructure: Cyclists’ perceptions of satisfaction and comfort. *Transportation Research Part A: Policy and Practice, 78*, 134-143. https://doi.org/10.1016/j.tra.2015.04.031

Carra, M., Maternini, G., & Barabino, B. (2022). On sustainable positioning of electric vehicle charging stations in cities: An integrated approach for the selection of indicators. *Sustainable Cities and Society, 85*, 104067. https://doi.org/10.1016/j.scs.2022.104067

Carrara, E., Ciavarella, R., Boglietti, S., Carra, M., Maternini, G., & Barabino, B. (2021). Identifying and selecting key sustainable parameters for the monitoring of e-powered micro personal mobility vehicles: Evidence from Italy. *Sustainability, 13*(16), 9226. https://doi.org/10.3390/su13169226

Castillo, H., & Pitfield, D. E. (2009). ELASTIC - A methodological framework for identifying and selecting sustainable transport indicators. *Transportation Research Part D: Transport and Environment, 15*(4), 179-188. https://doi.org/10.1016/j.trd.2009.09.002

European Committee for Standardization. (2002). *EN 13816: Public passenger transport - Service quality definition, targeting and measurement*. Brussels, Belgium: Author.

Chauhan, V., Gupta, A., & Parida, M. (2021). Demystifying service quality of multimodal transportation hub (MMTH) through measuring users' satisfaction of public transport. *Transport Policy, 102*, 47-60. https://doi.org/10.1016/j.tranpol.2021.01.004

Da Rocha, V. T., Brandli, L. L., Kalil, R. M. L., Salvia, A. L., & Prietto, P. D. M. (2019). Quality of sidewalks in a Brazilian city: A broad vision. *Theoretical and Empirical Researches in Urban Management, 14*(2), 41-58.

De Aquino Traldi, W., De França Marques, S., Pitombo, C. S., De Sousa, P. B., & De Melo, R. A. (2022). Avaliação da infraestrutura cicloviária e interpolação espacial de seus indicadores de qualidade: Uma abordagem baseada em análise hierárquica e geoestatística. *Revista Brasileira De Cartografia, 74*(4), 968-985. https://doi.org/10.14393/rbcv74n4-65916

De Oña, J., & De Oña, R. (2014). Quality of service in public transport based on customer satisfaction surveys: A review and assessment of methodological approaches. *Transportation Science, 49*(3), 605-622. https://doi.org/10.1287/trsc.2014.0544

Dell'Olio, L., Ibeas, A., & De Oña, J. (2017). Public transportation quality of service: Factors, models, and applications. https://international.scholarvox.com/catalog/book/docid/88849532

Dever, J. A., Rafferty, A., & Valliant, R. (2008). Internet surveys: Can statistical adjustments eliminate coverage bias? *Statistical Research Methods, 2*(2), 128. https://doi.org/10.18148/srm/2008.v2i2.128

Diana, M., Pirra, M., Castro, A., Duarte, A., Brangeon, V., Di Majo, C., Herrero, D., Hrin, G. R., & Woodcock, A. (2016). Development of an integrated set of indicators to measure the quality of the whole traveller experience. *Transportation Research Procedia, 14*, 1164-1173. https://doi.org/10.1016/j.trpro.2016.05.187

Diana, M., Duarte, A., & Pirra, M. (2017). Transport quality profiles of European cities based on a multidimensional set of satisfaction ratings indicators. *Transportation Research Record, 2643*(1), 84-92. https://doi.org/10.3141/2643-10

D'Orso, G., & Migliore, M. (2019). A GIS-based method for evaluating the walkability of a pedestrian environment and prioritised investments. *Journal of Transport Geography, 82*, 102555. https://doi.org/10.1016/j.jtrangeo.2019.102555

Doran, G. T. (1981). There's a SMART way to write management's goals and objectives. *Management Review, 70*, 35-36.

Douglas, N. J. (2015). Valuing public transport service quality using a combined rating & stated preference survey. In *Proceedings of the 37th Australasian Transport Research Forum (ATRF)*, Sydney, Australia, 30 September - 2 October 2015.

Du, Y., & Huang, W. (2022). Evaluation of street space quality using streetscape data: Perspective from recreational physical activity of the elderly. *ISPRS International Journal of Geo-Information, 11*(4), 241. https://doi.org/10.3390/ijgi11040241

Eboli, L., & Mazzulla, G. (2009). A new customer satisfaction index for evaluating transit service quality. *Journal of Public Transportation, 12*(3), 21-37. https://doi.org/10.5038/2375-0901.12.3.2

Eboli, L., & Mazzulla, G. (2010). A methodology for evaluating transit service quality based on subjective and objective measures from the passenger's point of view. *Transport Policy, 18*(1), 172-181. https://doi.org/10.1016/j.tranpol.2010.07.007

Ely, V. H. M. B., De Oliveira, J. M., & Logsdon, L. (2012). A bus stop shelter evaluated from the user's perspective. *Work, 41*, 1226-1233. https://doi.org/10.3233/wor-2012-0307-1226

European Committee for Standardization. (2006, April 25). *EN 15140:2006 Public passenger transport - Basic requirements and recommendations for systems that measure delivered service quality (CEN/TC 320)*. Brussels, Belgium: Author.

European 4th RTD Framework Programme. (1998). *Quality approach in tendering/contracting urban public transport operations (QUATTRO)*.

Ferretto, L., Carra, M., & Barabino, B. (2024). Key quality criteria in an integrated multiple transport systems scenario: A systematic literature review. In O. Gervasi, B. Murgante, C. Garau, D. Taniar, A. M. A. C. Rocha, & M. N. Faginas Lago (Eds.), *Computational science and its applications - ICCSA 2024 workshops* (Lecture Notes in Computer Science, Vol. 14823, pp. 114-132). Springer, Cham. https://doi.org/10.1007/978-3-031-65329-2_8

Fistola, R., Gallo, M., La Rocca, R. A., & Russo, F. (2020). The effectiveness of urban cycle lanes: From dyscrasias to potential solutions. *Sustainability, 12*(6), 2321. https://doi.org/10.3390/su12062321

Friman, M., & Fellesson, M. (2009). Service supply and customer satisfaction in public transportation: The quality paradox. *Journal of Public Transportation, 12*(4), 57-69. https://doi.org/10.5038/2375-0901.12.4.4

Fruin, J. J. (1971). Designing for pedestrians: A level-of-service concept. Paper presented at the 50th Annual Meeting of the Highway Research Board, Washington, DC.

Gössling, S. (2016). Urban transport justice. *Journal of Transport Geography, 54*, 1-9. https://doi.org/10.1016/j.jtrangeo.2016.05.002

Groenendijk, L., Rezaei, J., & Correia, G. (2018). Incorporating the travellers' experience value in assessing the quality of transit nodes: A Rotterdam case study. *Case Studies on Transport Policy, 6*(4), 564-576. https://doi.org/10.1016/j.cstp.2018.07.007

Hamerska, M., Ziółko, M., & Stawiarski, P. (2022). A sustainable transport system—The MMQUAL model of shared micromobility service quality assessment. *Sustainability, 14*(7), 4168. https://doi.org/10.3390/su14074168

Highway Research Board. (1965). *Highway capacity manual* (1965 ed.). National Research Council.

Hobbie, S. E., & Grimm, N. B. (2020). Nature-based approaches to managing climate change impacts in cities. *Philosophical Transactions of the Royal Society B: Biological Sciences, 375*(1794), 20190124. https://doi.org/10.1098/rstb.2019.0124

Hsu, T. P., & Lin, Y. T. (2011). A model for planning a bicycle network with multi-criteria suitability evaluation using GIS. *WIT Transactions on Ecology and the Environment, 1*, 243-252. https://doi.org/10.2495/rav110231

Hsu, C., Liou, J. J., Lo, H., & Wang, Y. (2018). Using a hybrid method for evaluating and improving the service quality of public bike-sharing systems. *Journal of Cleaner Production, 202*, 1131-1144. https://doi.org/10.1016/j.jclepro.2018.08.193

Hu, M., Lu, L., & Yang, J. (2019). Exploring an estimation approach for the pedestrian level of service for metro stations based on an interaction index. *Transportation Letters, 12*(6), 417-426. https://doi.org/10.1080/19427867.2019.1629061

Ismail, N. I. N., Rahman, N. A. A., Muhamad, N. S., Yaccob, A. A., & Mohtar, N. H. (2020). Pedestrian's perception toward quality of sidewalk facilities: Case study UiTM Pulau Pinang. *IOP Conference Series: Materials Science and Engineering, 849*(1), 012057. https://doi.org/10.1088/1757-899X/849/1/012057

Jahan, M. I., Mazumdar, A. A. B., Hadiuzzaman, M., & Mashrur, S. M. (2020). Analyzing service quality of pedestrian sidewalks under mixed traffic condition considering latent variables. *Journal of Urban Planning and Development, 146*(2), 04020011. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000587

Joumard, R., Gudmundsson, H., & Folkeson, L. (2011). Framework for assessing indicators of environmental impacts in the transport sector. *Transportation Research Record, 2242*(1), 55-63. https://doi.org/10.3141/2242-07

Julio, R., Monzon, A., & Susilo, Y. O. (2022). Identifying key elements for user satisfaction of bike-sharing systems: A combination of direct and indirect evaluations. *Transportation, 51*(2), 407-438. https://doi.org/10.1007/s11116-022-10335-3

Kazemzadeh, K., & Sprei, F. (2022). Towards an electric scooter level of service: A review and framework. *Travel Behaviour and Society, 29*, 149-164. https://doi.org/10.1016/j.tbs.2022.06.005

Kittleson & Associates, Inc., Texas Transportation Institute, & Transport Consulting Limited. (1999). *Transit capacity and quality of service manual* (1st ed.). Transportation Research Board, National Research Council.

Kittelson & Associates, Inc., Parsons Brinckerhoff, Inc., KFH Group, Inc., Texas A&M Transportation Institute, & Arup. (2013). *Transit capacity and quality of service manual* (3rd ed., TCRP Report No. 165). Transportation Research Board.

Larranaga, A. M., Arellana, J., Rizzi, L. I., Strambi, O., & Cybis, H. B. B. (2018). Using best-worst scaling to identify barriers to walkability: A study of Porto Alegre, Brazil. *Transportation, 46*(6), 2347-2379. https://doi.org/10.1007/s11116-018-9944-x

Liang, X., Chen, T., Ye, M., Lin, H., & Li, Z. (2021). A hybrid fuzzy BWM-VIKOR MCDM to evaluate the service level of bike-sharing companies: A case study from Chengdu, China. *Journal of Cleaner Production, 298*, 126759. https://doi.org/10.1016/j.jclepro.2021.126759

Liou, J. J., Hsu, C., & Chen, Y. (2014). Improving transportation service quality based on information fusion. *Transportation Research Part A: Policy and Practice, 67*, 225-239. https://doi.org/10.1016/j.tra.2014.07.007

Ma, F., Shi, W., Yuen, K. F., Sun, Q., & Guo, Y. (2019). Multi-stakeholders' assessment of bike sharing service quality based on DEMATEL-VIKOR method. *International Journal of Logistics Research and Applications, 22*(5), 449-472. https://doi.org/10.1080/13675567.2019.1568401

Marisamynathan, S., & Lakshmi, S. (2016). Method to determine pedestrian level of service for sidewalks in Indian context. *Transportation Letters, 10*(5), 294-301. https://doi.org/10.1080/19427867.2016.1264668

Meyer, M. D. (1995). Performance measures to improve transportation systems and agency operations: Report of a conference. Transportation Research Board.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moodley, S., & Venter, C. (2022). Measuring the service quality at multimodal public transport interchanges: A needs-driven approach. *Transportation Research Record, 2676*(10), 194-206. https://doi.org/10.1177/03611981221088782

National Academies of Sciences, Engineering, and Medicine. (2008). *Multimodal level of service analysis for urban streets*. Transportation Research Board. https://doi.org/10.17226/14175

Ojo, T. K. (2017). Quality of public transport service: An integrative review and research agenda. *Transportation Letters, 11*(2), 104-116. https://doi.org/10.1080/19427867.2017.1283835

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12-40.

Pattnaik, S. B., & Kumar, K. R. (1996). Level of service of urban roads based on users' perception. *Civil Engineering Systems, 14*(2), 87-110. https://doi.org/10.1080/02630259608970213

Phillips, R., & Guttenplan, M. (2003). A review of approaches for assessing multimodal quality of service. *Journal of Public Transportation, 6*(4), 69-87. https://doi.org/10.5038/2375-0901.6.4

Podciborski, T. (2017). A method for evaluating tram stops based on passenger expectations and the needs of disabled persons. In *Proceedings of 10th International Conference "Environmental Engineering"*. https://doi.org/10.3846/enviro.2017.115

Podgórniak-Krzykacz, A., Przywojska, J., & Trippner-Hrabi, J. (2022). A public value-based, multilevel evaluation framework to examine public bike-sharing systems: Implications for cities' sustainable transport policies. *Transport and Telecommunication Journal, 23*(2), 180-194. https://doi.org/10.2478/ttj-2022-0016

Pullen, W. T. (1993). Definition and measurement of quality of service for local public transport management. *Transport Reviews, 13*(3), 247-264. https://doi.org/10.1080/01441649308716849

Raicu, R., & Raicu, S. (2005). Complex aspects of transport quality. *WIT Transactions on the Built Environment, 77*. https://doi.org/10.2495/ut050281

Raswol, L. M. (2020). Qualitative assessment for walkability: Duhok University campus as a case study. *IOP Conference Series: Materials Science and Engineering, 978*(1), 012001. https://doi.org/10.1088/1757-899x/978/1/012001

Redman, L., Friman, M., Gärling, T., & Hartig, T. (2012). Quality attributes of public transport that attract car users: A research review. *Transport Policy, 25*, 119-127. https://doi.org/10.1016/j.tranpol.2012.11.005

Rodriguez-Valencia, A., Barrero, G. A., Ortiz-Ramirez, H. A., & Vallejo-Borda, J. A. (2020). Power of user perception on pedestrian quality of service. *Transportation Research Record, 2674*(5), 250-258. https://doi.org/10.1177/0361198120914611

Rosa, M. P., & Da Cruz Lopes, J. (2019). Senior tourists' perceptions of bus stop environments used in collaborative design. *WIT Transactions on the Built Environment.* https://doi.org/10.2495/ut190111

Rossetti, S. & Tiboni M. (2020). In field assessment of safety, security, comfort and accessibility of bus stops: A planning perspective. *European Transport / Trasporti Europei, 80* (ET.2020), 1-17. https://doi.org/10.48295/et.2020.80.8

Rowangould, G., & Corning-Padilla, A. (2019). Evaluating how the quality of pedestrian infrastructure affects the choice to walk. *LSU Scholarly Repository.* https://repository.lsu.edu/transet_pubs/60/

Russo, F., & Rindone, C. (2023). Logical framework approach in transportation planning: The passenger services in the Messina Strait. *Transportation Research Procedia, 69*, 855-862. https://doi.org/10.1016/j.trpro.2023.02.245

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology, 15*(3), 234-281. https://doi.org/10.1016/0022-2496(77)90033-5

Saaty, R. W. (1987). The analytic hierarchy process - What it is and how it is used. *Mathematical Modelling, 9*(3-5), 161-176. https://doi.org/10.1016/0270-0255(87)90473-8

Saaty, T. L. (1990). An exposition of the AHP in reply to the paper “remarks on the analytic hierarchy process.” *Management Science, 36*(3), 259-268.

Saaty, T. L. (1994). Highlights and critical points in the theory and application of the analytic hierarchy process. *European Journal of Operational Research, 74*(3), 426-447. https://doi.org/10.1016/0377-2217(94)90222-4

Sarkar, S. (2003). Qualitative evaluation of comfort needs in urban walkways in major activity centers. *Transportation Quarterly, 57.*

Shin, E. J. (2020). A comparative study of bike-sharing systems from a user’s perspective: An analysis of online reviews in three U.S. regions between 2010 and 2018. *International Journal of Sustainable Transportation, 15*(12), 908-923. https://doi.org/10.1080/15568318.2020.1830320

Soltani, A., Fazeli, S., & Eskandari, S. (2017). Comparing sidewalk design status from the pedestrians' perspective versus urban street design standards. *World Review of Intermodal Transportation Research, 6*(3), 193-208.

Sousa, A., Santos, B., & Goncalves, J. (2019). Pedestrian environment quality assessment in Portuguese medium-sized cities. *IOP Conference Series: Materials Science and Engineering, 471*, 062033. https://doi.org/10.1088/1757-899X/471/6/062033

Talavera-Garcia, R., & Soria-Lara, J. A. (2015). Q-PLOS, developing an alternative walking index: A method based on urban design quality. *Cities, 45*, 7-17. https://doi.org/10.1016/j.cities.2015.03.003

Tan, D., Wang, W., Lu, J., & Bian, Y. (2007). Research on methods of assessing pedestrian level of service for sidewalk. *Journal of Transportation Systems Engineering and Information Technology, 7*(5), 74-79. https://doi.org/10.1016/s1570-6672(07)60041-5

Transportation Research Board. (1997). Real-time human perceptions: Toward a bicycle level of service. *Transportation Research Record, 1578*, 119-126. https://doi.org/10.3141/1578-15

Transportation Research Board. (2003). *Travel demand and land use 2003: Planning and administration.* Transportation Research Record, No. 1831. National Research Council.

Ujjwal, J., & Bandyopadhyaya, R. (2021). Development of pedestrian level of service assessment guidelines for mixed land use areas considering quality of service parameters. *Transportation in Developing Economies, 7*(1). https://doi.org/10.1007/s40890-021-00113-8

Ujjwal, L., & Bandyopadhyaya, R. (2023). Development of pedestrian level of service (PLOS) model and satisfaction perception rating models for pedestrian infrastructure for mixed land-use urban areas. *Transportation, 50*(2), 355-381. https://doi.org/10.1007/s11116-021-10247-8

United States Agency for International Development. (1969). *Logical framework: A planning and evaluation tool.* United States Agency for International Development.

Vallejo-Borda, J. A., Cantillo, V., & Rodriguez-Valencia, A. (2020). A perception-based cognitive map of the pedestrian perceived quality of service on urban sidewalks. *Transportation Research Part F: Traffic Psychology and Behaviour, 73*, 107-118. https://doi.org/10.1016/j.trf.2020.06.013

Vujičić, M., & Prester, J. (2019). Assessing service quality of public tram transport in Zagreb city using P-TRANSQUAL mode. *Zbornik Ekonomskog Fakulteta u Zagrebu, 17*(1), 19–31. https://doi.org/10.22598/zefzg.2019.1.19

Xue, X., Wang, Z., Liu, X., Zhou, Z., & Song, R. (2022). A choice behavior model of bike-sharing based on user perception, psychological expectations, and loyalty. *Journal of Advanced Transportation, 2022*, 1-14. https://doi.org/10.1155/2022/6695977

Wang, T., & Lee, H. (2008). Developing a fuzzy TOPSIS approach based on subjective weights and objective weights. *Expert Systems with Applications, 36*(5), 8980-8985. https://doi.org/10.1016/j.eswa.2008.11.035

Wind, Y., & Saaty, T. L. (1980). Marketing applications of the analytic hierarchy process. *Management Science, 26*(7), 641-658. https://doi.org/10.1287/mnsc.26.7.641

Zelany, M. (1974). A concept of compromise solutions and the method of the displaced ideal. *Computers & Operations Research, 1*(3–4), 479-496. https://doi.org/10.1016/0305-0548(74)90064-1

7.1 Publications arising from this research

Ferretto, L., Carra, M., & Barabino, B. (2023, June 15-16). Mobilità non motorizzata: Una revisione sistematica della letteratura sui principali parametri di qualità. In *Atti della XXV Conferenza SIU 2023* (Cagliari, Italy).

Ferretto, L., Carra, M., & Barabino, B. (2024). Key quality criteria in an integrated multiple transport systems scenario: A systematic literature review. In O. Gervasi, B. Murgante, C. Garau, D. Taniar, A. M. A. C. Rocha, & M. N. Faginas Lago (Eds.), *Computational Science*

and Its Applications – ICCSA 2024 Workshops. ICCSA 2024. Lecture Notes in Computer Science, 14823 (pp. 114-132). Springer, Cham. https://doi.org/10.1007/978-3-031-65329-2_8. Winner of ICCSA 2024 Best Paper Award.

Ferretto, L., Barabino, B., & Currie, G. (2025). Assessing interaction quality between multimodal transit interchanges: A Melbourne case study. In *Australasian Transport Research Forum 2025 Proceedings* (18-21 November, Auckland, New Zealand).

Ferretto, L., Carra, M., & Barabino, B. (2026). Enlarging the concept of quality of service: from a sectoral to a multimodal perspective. Transportation Letters, 1–16. https://doi.org/10.1080/19427867.2026.2636006. Winner of the Best Younger Researcher award for the best paper at the *XXVII International Conference LWC 2025, Brescia (Italy).*

Ferretto, L., Carra, M., & Barabino, B. (2026). Key quality parameters (KQPs) driving the development of sustainable multimodal transport. *Transportation Research Interdisciplinary Perspectives.* Accepted, vol.36, ISSN 2590-1982, https://doi.org/10.1016/j.trip.2026.101924.

8. APPENDIX

8.1 Appendix A

Day: 13/09/2025
Hour: 08:15

O-D: 3. Valdonega - Cittadella
path: number 1
segment: number 3

Criteria	Parameter I	Parameter II	Indicator	not present	present
availability	network	areas covered	LPT service/infrastructure		x
availability	network	areas covered	pedestrian infrastructure		x
availability	network	areas covered	cycling infrastructure	x	
availability	network	areas covered	bike-sharing service		x
availability	network	areas covered	scooter-sharing service	x	

A. 1. Excerpt from the technical checklist used during the Mystery Client field measurements to evaluate delivered quality.

8.2 Appendix B

ID 1	Methodological Quality	Quality Relevance
Methodological Quality	1,0000	1,0000
Quality Relevance	1,0000	1,0000

ID 2	Methodological Quality	Quality Relevance
Methodological Quality	1,0000	5,0000
Quality Relevance	0,2000	1,0000

ID 3	Methodological Quality	Quality Relevance
Methodological Quality	1,0000	1,0000
Quality Relevance	1,0000	1,0000

ID 4	Methodological Quality	Quality Relevance
Methodological Quality	1,0000	1,0000
Quality Relevance	1,0000	1,0000

Figure B.1. Excerpt from the Excel matrix structured to disaggregate the academic experts' ratings, where each ID number corresponds to one academic expert. In the original file, this comparison of the two components includes 27 matrices corresponding to the 27 academic experts' respondents.

ID 1	Measurability	Easy Availability	Speed Availability	Interpretability	Relevance to multiple t.s.
Measurability	1,0000	0,2500	9,0000	6,0000	9,0000
Easy Availability	4,0000	1,0000	6,0000	4,0000	9,0000
Speed Availability	0,1111	0,1667	1,0000	0,5000	4,0000
Interpretability	0,1667	0,2500	2,0000	1,0000	4,0000
Relevance to multiple transport systems	0,1111	0,1111	0,2500	0,2500	1,0000

ID 2	Measurability	Easy Availability	Speed Availability	Interpretability	Relevance to multiple t.s.
Measurability	1,0000	3,0000	3,0000	3,0000	4,0000
Easy Availability	0,3333	1,0000	4,0000	0,3333	0,5000
Speed Availability	0,3333	0,2500	1,0000	0,3333	2,0000
Interpretability	0,3333	3,0000	3,0000	1,0000	3,0000
Relevance to multiple transport systems	0,2500	2,0000	0,5000	0,3333	1,0000

Figure B.2. Excerpt from the Excel matrix structured to disaggregate the academic experts' ratings, where each ID number corresponds to one academic expert. In the original file, this comparison of the two components includes 24 matrices corresponding to the 27 academic experts' respondents minus the three evaluations excluded because they lacked credibility.

8.3 Appendix C

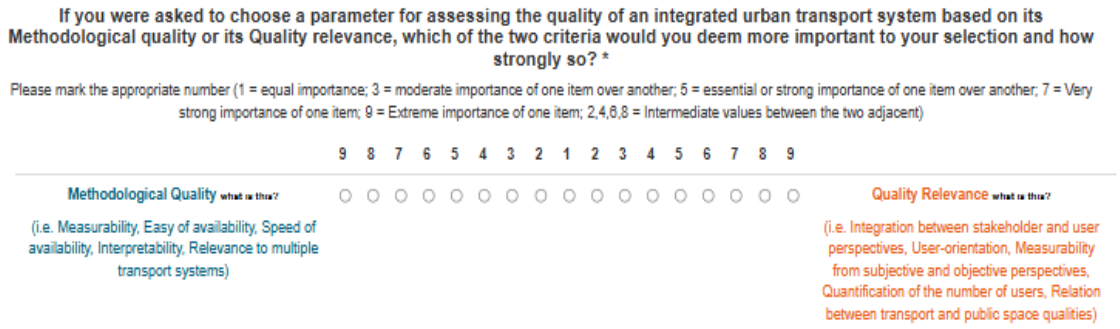


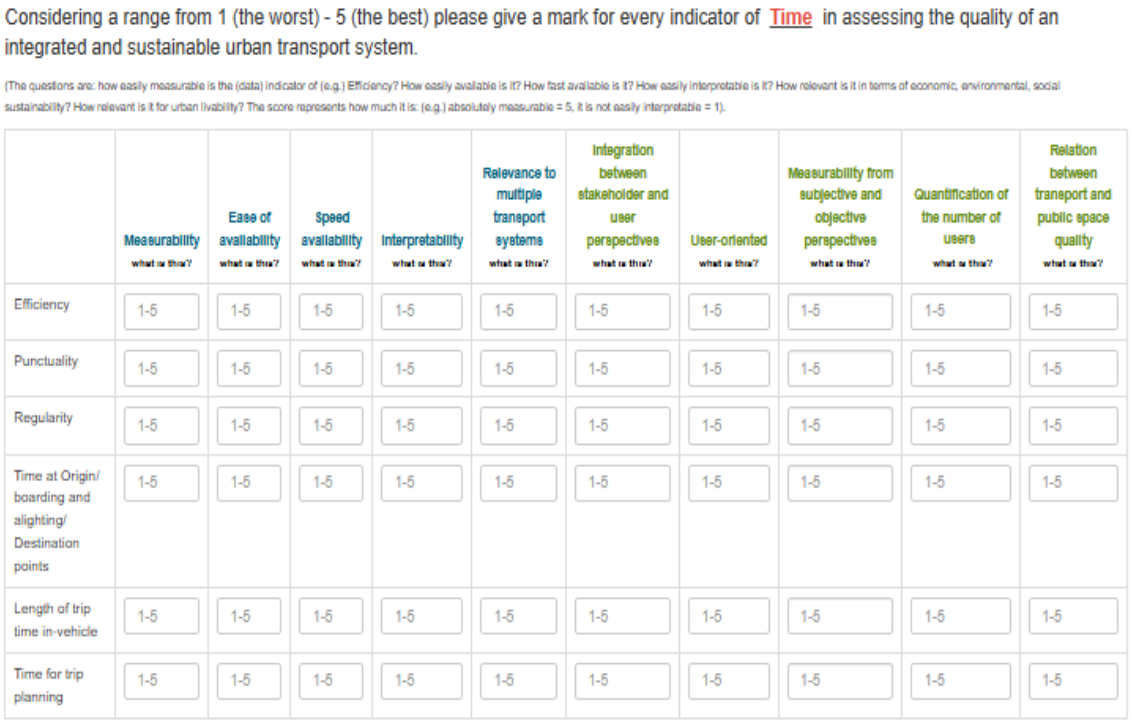
Figure C.1. Excerpt from the structured web survey comparing the two components.

In deciding an indicator of integrated quality of an urban transport system, which of the following attributes would you deem more important for indicator choice and how strongly? *

Please mark the appropriate number (1 = equal importance; 3 = moderate importance of one item over another; 5 = essential or strong importance of one item over another; 7 = Very strong importance of one item; 9 = Extreme importance of one item; 2,4,6,8 = Intermediate values between the two adjacent)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Measurability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Easy Availability what is this?
Easy Availability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Speed Availability what is this?
Speed Availability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Interpretability what is this?
Interpretability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Relevance to multiple transport systems what is this?
Relevance to multiple transport systems what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Measurability what is this?
Measurability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Speed Availability what is this?
Speed Availability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Relevance to multiple transport systems what is this?
Relevance to multiple transport systems what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Easy Availability what is this?
Easy Availability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Interpretability what is this?
Interpretability what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Measurability what is this?
Integration between stakeholder and user perspectives what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	User-orientation what is this?
User-orientation what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Measurability from subjective and objective perspectives what is this?
Measurability from subjective and objective perspectives what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Quantification of the number of users what is this?
Quantification of the number of users what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Relation between transport and public space qualities what is this?
Relation between transport and public space qualities what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Integration between stakeholder and user perspectives what is this?
Integration between stakeholder and user perspectives what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Measurability from subjective and objective perspectives what is this?
Measurability from subjective and objective perspectives what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Relation between transport and public space qualities what is this?
Relation between transport and public space qualities what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	User-orientation what is this?
User-orientation what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Quantification of the number of users what is this?
Quantification of the number of users what is this?	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	Integration between stakeholder and user perspectives what is this?

Figure C.2. Excerpt from the structured web survey comparing the ten attributes.


Considering a range from 1 (the worst) - 5 (the best) please give a mark for every indicator of Time in assessing the quality of an integrated and sustainable urban transport system.

(The questions are: how easily measurable is the (data) indicator of (e.g.) Efficiency? How easily available is it? How fast available is it? How easily interpretable is it? How relevant is it in terms of economic, environmental, social sustainability? How relevant is it for urban livability? The score represents how much it is: (e.g.) absolutely measurable = 5, it is not easily interpretable = 1).

	Measurability what is this?	Ease of availability what is this?	Speed availability what is this?	Interpretability what is this?	Relevance to multiple transport systems what is this?	Integration between stakeholder and user perspectives what is this?	User-oriented what is this?	Measurability from subjective and objective perspectives what is this?	Quantification of the number of users what is this?	Relation between transport and public space quality what is this?
Efficiency	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5
Punctuality	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5
Regularity	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5
Time at Origin/ boarding and alighting/ Destination points	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5
Length of trip time in-vehicle	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5
Time for trip planning	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5	1-5

Figure C.3. Excerpt from the structured web survey comparing the 65 quality parameters with attributes. In particular, this figure shows the structure for evaluating the criterion of Time.

8.4 Appendix D

Table D.1. Results of the multi-criteria weighting process: individual expert weights for each indicator and corresponding aggregated statistics (mean, standard deviation, coefficient of variation) used to validate weight consistency. Indicators that do not require weighting - because they singularly represent an entire parameter and are therefore not comparable with others - were not submitted to expert judgment; instead, they were assigned the maximum score (1000). These include Customer Service, Access to a Hotline, and Motorized Flow.

Indicators	weight of E1	weight of E2	weight of E3	weight of E4	mean	st. dev.	coeff. of variation
Accessibility / Internal interface / Entrance/exits							
Ramps at LPT stops	0.333	0.105	0.163	0.540	0.285	0.195	0.685
Ramps to get on and off to the pedestrian infrastructure	0.333	0.258	0.297	0.163	0.263	0.073	0.278
Ramps to enter and exit to the cycling infrastructure	0.333	0.637	0.540	0.297	0.452	0.163	0.362
Availability / Network / Area covered							
LPT service/infrastructure	0.152	0.148	0.159	0.228	0.172	0.038	0.219
Pedestrian infrastructure	0.205	0.199	0.237	0.169	0.203	0.028	0.136
Cycling infrastructure	0.132	0.136	0.131	0.153	0.138	0.010	0.075
Bike-sharing service	0.083	0.077	0.077	0.088	0.081	0.005	0.066
Scooters-sharing service	0.054	0.038	0.038	0.054	0.056	0.016	0.292
Continuity of the pedestrian infrastructure	0.205	0.233	0.223	0.153	0.195	0.030	0.152
Continuity of the cycling infrastructure	0.168	0.136	0.136	0.153	0.155	0.014	0.092
Comfort / Ambient conditions / Cleanliness							
Bins at the LPT stops	0.093	0.093	0.093	0.133	0.103	0.020	0.194
Bins along the pedestrian infrastructure	0.098	0.098	0.093	0.118	0.101	0.011	0.108
Trash/broken glasses at the LPT stops	0.204	0.204	0.215	0.198	0.205	0.007	0.036
Trash/broken glasses along the pedestrian infrastructure	0.215	0.215	0.215	0.181	0.206	0.017	0.081
Trash/broken glasses along the cycling infrastructure	0.215	0.215	0.215	0.181	0.206	0.017	0.081
Illegal graffiti at the LPT stops	0.066	0.066	0.060	0.086	0.069	0.011	0.164
Illegal graffiti along the pedestrian infrastructure	0.051	0.051	0.060	0.052	0.053	0.005	0.089

Illegal graffiti along the cycling infrastructure	0.060	0.060	0.047	0.052	0.055	0.007	0.123
Comfort / Complementary facilities / Amenities							
Seating facility at the LPT stops	0.232	0.234	0.239	0.158	0.216	0.039	0.179
Seating facility along the pedestrian infrastructure	0.088	0.089	0.130	0.135	0.111	0.025	0.229
Public toilet	0.156	0.157	0.086	0.165	0.141	0.037	0.262
Bicycle racks at interchanges	0.210	0.200	0.239	0.111	0.190	0.055	0.291
Bicycle racks along the route	0.094	0.100	0.130	0.062	0.096	0.028	0.290
Drinking water facility	0.110	0.110	0.106	0.193	0.130	0.042	0.234
Number of water facility	0.110	0.110	0.070	0.177	0.117	0.044	0.379
Customer care / Assistance / Customer orientation							
Customer service	/	/	/	/	/	/	/
Customer care / Commitment / Innovation and initiatives							
Access to a hotline	/	/	/	/	/	/	/
Environmental impact / Pollution and health / Natural resources							
Land consumption // natural surfaces	0.333	0.333	0.333	0.500	0.375	0.083	0.222
Amount of natural surfaces (m^2)	0.667	0.667	0.667	0.500	0.625	0.083	0.133
Information / General information / Info. about accessibility							
Signs about the access to the LPT infrastructure/service	0.260	0.260	0.261	0.176	0.239	0.042	0.176
Signs about the access to the pedestrian infrastructure	0.197	0.197	0.172	0.099	0.166	0.047	0.280
Signs about the access to the cycling infrastructure	0.149	0.149	0.172	0.099	0.142	0.031	0.218
Tactile information at nodal points (intersections, major crossings)	0.197	0.197	0.198	0.313	0.226	0.058	0.257
Information about disabled friendly interchanges	0.197	0.197	0.198	0.313	0.226	0.058	0.257
Safety & Security / Freedom from accidents / Traffic flow							
Motorized flow (number of motorized vehicle per hour)	/	/	/	/	/	/	/
Safety & Security / Freedom from crime / Lighting							
Working lighting system at the LPT stops	0.143	0.111	0.163	0.540	0.239	0.201	0.842
Working lighting system along the pedestrian infrastructure	0.429	0.444	0.540	0.297	0.427	0.100	0.234
Working lighting system along the cycling infrastructure	0.429	0.444	0.297	0.163	0.333	0.131	0.393
Time / Length of trip time / At Origin/boarding & alighting/Destination points							
Min/h spent waiting the LPT	0.667	0.667	0.750	0.667	0.688	0.042	0.061
Min/h spent looking for a shared micro-mobility means of transport	0.333	0.333	0.250	0.333	0.313	0.042	0.133
Time / Length of the trip time / Transfer time							
Min/h spent to reach the destination with the LPT	0.540	0.493	0.540	0.540	0.528	0.023	0.044
Min/h spent to reach the destination by walking	0.163	0.196	0.163	0.297	0.205	0.063	0.309
Min/h spent to reach the destination by cycling	0.297	0.311	0.297	0.163	0.267	0.069	0.260

Table D.2 . Example of the Mystery Client measurement dataset structure for O-D 1. This table illustrates how raw field observations are translated into the delivered quality scores used in the analysis. For each indicator, the reference measurement scale (presence - p/not presence - np), the corresponding numerical value, and the results of the hierarchical aggregation process (segment → path → O-D level) are reported. The example refers to O-D 1 and is representative of the structure adopted for all O-D pairs assessed through the Mystery Client technique.

Borgo Trento - Città antica	**1**				**2**			**3**		
	1	**2**	**3**	**4**	**1**	**2**	**3**	**1**	**2**	**3**
ramps at LPT stops	p									
ramps to get on and off to the pedestrian infrastructure	p	p	p	p	p	p	p	p	p	p
ramps to enter and exit to the cycling infrastructure	np	np	np	np	np	np	np	np	np	np
LPT service/infrastructure	p									
pedestrian infrastructure	p	p	p	p	p	p	p	p	p	p
cycling infrastructure	np	np	np	np	np	np	np	np	np	np
bike-sharing service	p	p	np	np	np	p	np	np	p	np
scooters-sharing service	p	p	p	p	p	p	p	p	p	p
continuity of the pedestrian infrastructure	p	p	p	p	p	p	p	p	p	p

continuity of the cycle infrastructure	np	np	np	np	np	np	np	np	np	np
bins at the LPT stops	p									
bins along the pedestrian infrastructure	p	p	p	p	p	p	p	p	p	p
trash/broken glasses at the LPT stops	p									
trash/broken glasses along the pedestrian infrastructure	p	np	p	p	p	p	p	p	np	p
trash/broken glasses along the cycling infrastructure	np	np	np	np	np	np	np	np	np	np
illegal graffiti at the LPT stops	p									
illegal graffiti along the pedestrian infrastructure	np	np	np	p	p	p	p	np	p	np
illegal graffiti along the cycling infrastructure	np	np	np	np	np	np	np	np	np	np
seating facility at the LPT stops (considering the boarding stop)	np									
seating facility along the pedestrian infrastructure	p	p	np	np	p	p	p	np	p	np
public toilet	np	np	np	np	np	np	np	np	np	np
bicycle racks at interchanges	p	p	np	np	p	p	p	p	p	p
bicycle racks along the route	p	p	np	np	p	p	p	p	p	p
drinking water facility	p	np	np	np	np	np	p	np	np	np
number of drinking water facility	1	0	0	0	0	0	1	0	0	0
customer service	np									
access to a hotline	np	np	np	p	np	np	p	np	np	p
land consumption // natural surfaces	p	np	np	np	np	p	np	np	p	np
amount of natural surfaces (m^2)	5423	0	0	0	0	11248	0	0	717	0
signs about the access to the LPT infrastructure/service	np									
signs about the access to the pedestrian infrastructure	p	p	p	p	np	np	np	np	p	np
signs about the access to the cycling infrastructure	np	np	np	np	np	np	np	np	np	np
provision of tactile information at nodal points (intersections, major crossings)	np	np	np	np	np	np	np	np	np	np
information about disabled friendly interchanges	np	np	np	np	np	p	np	np	np	np
motorized flow (number of motorized vehicle per hour)	340	680	340	0	400	400	0	400	800	0
working lighting system at the LPT stops	p									
working lighting system along the pedestrian infrastructure	p	p	p	p	p	p	p	np	p	P
working lighting system along the cycling infrastructure	np	np	np	np	np	np	np	np	np	np
min/h spent waiting the LPT	4									
min/h spent looking for a shared micro-mobility means of transport	>5	<5	<5	<5	>5	<5	<5	>5	<5	<5
min/h spent to reach the destination with the LPT	13									
min/h spent to reach the destination by walking	19:12				17:17			17:40		
min/h spent to reach the destination by cycling	06:24				05:46			05:53		
Borgo Trento - Città antica	**Average scores at segment level**									
	1				**2**			**3**		
	1	**2**	**3**	**4**	**1**	**2**	**3**	**1**	**2**	**3**
ramps at LPT stops	285									
ramps to get on and off to the pedestrian infrastructure	263	263	263	263	263	263	263	263	263	263
ramps to enter and exit to the cycling infrastructure	0	0	0	0	0	0	0	0	0	0
LPT service/infrastructure	172									
pedestrian infrastructure	203	203	203	203	203	203	203	203	203	203
cycling infrastructure	0	0	0	0	0	0	0	0	0	0
bike-sharing service	81	81	0	0	0	81	0	0	81	0
scooters-sharing service	56	56	56	56	56	56	56	56	56	56
continuity of the pedestrian infrastructure	195	195	195	195	195	195	195	195	195	195
continuity of the cycle infrastructure	0	0	0	0	0	0	0	0	0	0

bins at the LPT stops	103									
bins along the pedestrian infrastructure	101	101	101	101	101	101	101	101	101	101
trash/broken glasses at the LPT stops	205									
trash/broken glasses along the pedestrian infrastructure	206	0	206	206	206	206	206	206	0	206
trash/broken glasses along the cycling infrastructure	0	0	0	0	0	0	0	0	0	0
illegal graffiti at the LPT stops	69									
illegal graffiti along the pedestrian infrastructure	0	0	0	53	53	53	53	0	53	0
illegal graffiti along the cycling infrastructure	0	0	0	0	0	0	0	0	0	0
seating facility at the LPT stops (considering the boarding stop)	0									
seating facility along the pedestrian infrastructure	111	111	0	0	111	111	111	0	111	0
public toilet	0	0	0	0	0	0	0	0	0	0
bicycle racks at interchanges	190	190	0	0	190	190	190	190	190	190
bicycle racks along the route	96	96	0	0	96	96	96	96	96	96
drinking water facility	130	0	0	0	0	0	130	0	0	0
number of drinking water facility	0	0	0	0	0	0	0	0	0	0
customer service	0									
access to a hotline	0	0	0	1000	0	0	1000	0	0	1000
land consumption // natural surfaces	375	0	0	0	0	375	0	0	375	0
amount of natural surfaces (m^2)	0	0	0	0	0	625	0	0	0	0
signs about the access to the LPT infrastructure/service	0									
signs about the access to the pedestrian infrastructure	166	166	166	166	0	0	0	0	166	0
signs about the access to the cycling infrastructure	0	0	0	0	0	0	0	0	0	0
provision of tactile information at nodal points (intersections, major crossings)	0	0	0	0	0	0	0	0	0	0
information about disabled friendly interchanges	0	0	0	0	0	226	0	0	0	0
motorized flow (number of motorized vehicle per hour)	0	0	0	1000	0	0	1000	0	0	1000
working lighting system at the LPT stops	239									
working lighting system along the pedestrian infrastructure	427	427	427	427	427	427	427	0	427	427
working lighting system along the cycling infrastructure	0	0	0	0	0	0	0	0	0	0
min/h spent waiting the LPT	688									
min/h spent looking for a shared micro-mobility means of transport	0	313	313	313	0	313	313	0	313	313
min/h spent to reach the destination with the LPT	528									
min/h spent to reach the destination by walking	205			205			205			
min/h spent to reach the destination by cycling	267			267			267			
Borgo Trento - Città antica	**Average scores at path level**									
	1			**2**			**3**			
ramps at LPT stops	285									
ramps to get on and off to the pedestrian infrastructure	263			263			263			
ramps to enter and exit to the cycling infrastructure	0			0			0			
LPT service/infrastructure	172									
pedestrian infrastructure	203			203			203			
cycling infrastructure	0			0			0			
bike-sharing service	40.5			27.0			27.0			
scooters-sharing service	56			56			56			
continuity of the pedestrian infrastructure	195			195			195			
continuity of the cycle infrastructure	0			0			0			
bins at the LPT stops	103									

bins along the pedestrian infrastructure	101	101	101
trash/broken glasses at the LPT stops	205		
trash/broken glasses along the pedestrian infrastructure	154.5	206	137.3
trash/broken glasses along the cycling infrastructure	0	0	0
illegal graffiti at the LPT stops	69		
illegal graffiti along the pedestrian infrastructure	13.3	53	17.7
illegal graffiti along the cycling infrastructure	0	0	0
seating facility at the LPT stops (considering the boarding stop)	0		
seating facility along the pedestrian infrastructure	55.5	111	37
public toilet	0	0	0
bicycle racks at interchanges	95	190	190
bicycle racks along the route	48	96	96
drinking water facility	32.5	43.3	0
number of drinking water facility	0	0	0
customer service	0		
access to a hotline	250	333.3	333.3
land consumption // natural surfaces	93.8	125	125
amount of natural surfaces (m^2)	0	208.3	0
signs about the access to the LPT infrastructure/service	0		
signs about the access to the pedestrian infrastructure	83	0	55.3
signs about the access to the cycling infrastructure	0	0	0
provision of tactile information at nodal points (intersections, major crossings)	0	0	0
information about disabled friendly interchanges	0	75.3	0
motorized flow (number of motorized vehicle per hour)	250	333.3	333.3
working lighting system at the LPT stops	239		
working lighting system along the pedestrian infrastructure	427	427	284.7
working lighting system along the cycling infrastructure	0	0	0
min/h spent waiting the LPT	688		
min/h spent looking for a shared micro-mobility means of transport	234.8	208.7	208.7
min/h spent to reach the destination with the LPT	528		
min/h spent to reach the destination by walking	205	205	205
min/h spent to reach the destination by cycling	267	267	267
Borgo Trento - Città antica	**Average scores at O-D level**		
ramps at LPT stops	285		
ramps to get on and off to the pedestrian infrastructure	263		
ramps to enter and exit to the cycling infrastructure	0		
LPT service/infrastructure	172		
pedestrian infrastructure	203		
cycling infrastructure	0		
bike-sharing service	31.5		
scooters-sharing service	56		
continuity of the pedestrian infrastructure	195		
continuity of the cycle infrastructure	0		
bins at the LPT stops	103		
bins along the pedestrian infrastructure	101		

trash/broken glasses at the LPT stops	205
trash/broken glasses along the pedestrian infrastructure	169.5
trash/broken glasses along the cycling infrastructure	0
illegal graffiti at the LPT stops	69
illegal graffiti along the pedestrian infrastructure	28
illegal graffiti along the cycling infrastructure	0
seating facility at the LPT stops (considering the boarding stop)	0
seating facility along the pedestrian infrastructure	67.8
public toilet	0
bicycle racks at interchanges	158.3
bicycle racks along the route	80
drinking water facility	25.3
number of drinking water facility	0
customer service	0
access to a hotline	305.6
land consumption // natural surfaces	114.6
amount of natural surfaces (m^2)	69.4
signs about the access to the LPT infrastructure/service	0
signs about the access to the pedestrian infrastructure	46.1
signs about the access to the cycling infrastructure	0
provision of tactile information at nodal points (intersections, major crossings)	0
information about disabled friendly interchanges	25.1
motorized flow (number of motorized vehicle per hour)	305.6
working lighting system at the LPT stops	239
working lighting system along the pedestrian infrastructure	379.6
working lighting system along the cycling infrastructure	0
min/h spent waiting the LPT	688
min/h spent looking for a shared micro-mobility means of transport	217.4
min/h spent to reach the destination with the LPT	528
min/h spent to reach the destination by walking	205
min/h spent to reach the destination by cycling	267
All indicators	**5,599**

8.5 Appendix E

Table E.1. List of variables that make up the formulas in Chapter 4 of Methodology.

Symbol	Definition/Description	Formulas
S	the set of components or attributes	(1) (2) (3) (4) (5)
J	the set of academic experts involved	
$v_{sq,j}$	the numerical judgment provided by an academic expert j ∈ J of the pairwise comparison between a pair of components or attributes s ∈ S and q ∈ S	
V_{sj}	the overall un-normalised weight of component or attribute s ∈ S returned by academic expert j ∈ J	
$CI_{(j)}$	the consistency index and assesses whether the judgments provided by academic expert j ∈ J are consistent with the choices reported in the survey	
$\lambda_{\max\,(j)}$	the maximum eigenvalue required for computing the measure of consistency by academic expert j ∈ J	

RCI	be the random consistency index, which represents a CI function tabulated based on the maximum number of quality components or attributes (Saaty, 1980).	
P	the set of KQPs	(6) (7) (8) (9) (10) (11) (12)
H	the set of the methodological attributes	
K	the set of the quality attributes	
m_j	the weight of the methodological quality component as derived by AHP, according to the evaluation of academic expert j ∈ J	
q_j	the weight of the quality relevance component as derived by AHP, according to the evaluation of academic expert j ∈ J	
w_{jh}	the weight of attribute h ∈ H as derived by AHP, according to the evaluation of academic expert j ∈ J	
w_{jk}	the weight of attribute k ∈ K as derived by AHP, according to the evaluation of academic expert j ∈ J	
$\underline{M}$	the average weight of the methodological quality component	
$\underline{Q}$	the average weight of the quality relevance component	
$\underline{w_h}$	the average weight of attribute h ∈ H	
$\underline{w_k}$	the average weight of attribute k ∈ K	
V_{pjh}	the mark of parameter p ∈ P for attribute h ∈ H according to the evaluation of academic expert j ∈ J	
V_{pjk}	the mark of parameter p ∈ P for attribute k ∈ K according to the evaluation of academic expert j ∈ J	
V_{ph}	the average mark of parameter p ∈ P for attribute h ∈ H	
$\underline{V_{pk}}$	the average mark of parameter p ∈ P for attribute k ∈ K	
c	the quality criterion (e.g., Accessibility, Comfort, Safety & Security)	(13) (14) (15) (16) (17)
$P(c)$	the set of KQPs belonging to criterion c	
$I(p)$	the set of indicators belonging to KQP p	
$T(o,d)$	the set of path belonging to O-D pair o,d	
$L(t)$	the set of segments belonging to path $t \in T(o,d)$	
τ	the conformity threshold corresponding to the designed quality objective (e.g.,0.8; 800; 80%)	
x_p^l	the binary indicator of segment l for KQP p (=1 if the segment $l \in L(t)$ satisfies the quality standard for KQP p; 0 otherwise)	
U_l	the number of users travelling on segment $l \in L(t)$	
W_p	the weight of KQP p within criterion c	
W_i	the weight of indicator i within KQP p	
obs_l	the observation associated with the segment l	
$Rate_p^{obs}$	the $Rate$ of the observation of the segment l for each KQP p	
PUC_p^l	the Percentage of Users moving along the segment l for which a KQP p is compliant	
PUC_p^t	the Percentage of Users moving along the path t for which a KQP p is compliant	
$\overline{PUC_p^t}$	the average Percentage of Users moving along the path t for which a KQP p is compliant	
c	the quality criterion (e.g., Accessibility, Comfort, Safety & Security)	(18)
$P(c)$	the set of KQPs belonging to criterion c	
$I(p)$	the set of indicators belonging to KQP p	
S_{pi}	the score observed for indicator i belonging to KQP p	
W_{pi}	the corresponding weight	
U	the set of users	(19)
GI	the judgment of importance	
$GI_{t,p}^{(u)}$	the judgment of importance of the KQP p given by the $u \in U$ who moves using the using t ∈ $T(o,d)$	
$PUI_{t,p}$	the Percentage of Users U moving along the path t ∈ $T(o,d)$ for whom a KQP p is important	
U	the set of users	(20)
GS	the judgment of satisfaction	

$GS_{t,p}^{(u)}$	the judgment of satisfaction of the KQP p given by the $u \in U$ who moves using the using t ∈ $T(o,d)$	
$PUS_{t,p}$	the Percentage of Users U moving along the path t ∈ $T(o,d)$ who are satisfied about a KQP p	
PUT	the percentage of users targeted	(21) (22) (23) (24)
PUI	the percentage of users for whom the parameter is important	
PUC	the percentage of users in conformity	
PUS	the percentage of users satisfied	
$QG1$	the difference between the percentage of users targeted and the percentage of users for whom the parameter is important	
$QG2$	the difference between the percentage of users in conformity and the percentage of users targeted	
$QG3$	the difference between the percentage of users satisfied and the percentage of users in conformity	
$QG4$	the difference between the percentage of users satisfied and the percentage of users for whom the parameter is important	