International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(15s) 2026 pp. 501-516
https://doi.org/10.70917/ijcisim-2026-4418
Article

Design of a strategic model for restaurants and bars based on continuous quality service improvement, Barranco 2025.

Mario Enrique Gerardo Gastelú Carbajal

Universidad Cesar Vallejo, Perú., egasteluca@ucvvirtual.edu.pe

Abstract: Service quality bears directly on organisational sustainability, yet most SERVQUAL research captures the customer perspective, leaving the provider viewpoint under-examined in gastronomy. This study assessed how service-providing staff perceive the quality they deliver across five SERVQUAL dimensions and four establishment types in Barranco, Lima, Peru. Ninety informants from 30 food-and-beverage establishments completed an adapted 10-item SERVQUAL instrument. Kruskal–Wallis tests, Dunn–Bonferroni post-hoc comparisons, effect sizes (ε^2) and exploratory factor analysis were applied. All six comparisons yielded statistically significant differences with large effect sizes (ε^2 range: 0.147–0.363). Responsiveness was the weakest dimension (M = 3.30; 26.67% favourable) and tangibles the strongest (M = 4.16; 72.22% favourable). Fine dining showed the strongest overall quality profile. The resulting deficit hierarchy provides an empirical basis for establishment-specific continuous improvement strategies within Barranco's gastronomic sector.

Keywords: service quality; SERVQUAL; provider perspective; Gap 1; food and beverage; Barranco; Lima; Peru; Kruskal–Wallis; establishment type; continuous improvement; hospitality management; restaurant quality; gastronomy; developing economy.

1. Introduction

Service quality is a central concern in business management, with a direct bearing on customer satisfaction and long-term organisational viability (Cayo-Velásquez et al., 2025). In the food-and-beverage sector of developing economies, where informal operations, high staff turnover and absent quality frameworks predominate, the distance between what providers intend to deliver and what the customer actually experiences can be considerable (Olawuyi and Kleynhans, 2025). Peru's gastronomy has gained international stature — Lima features prominently in global culinary rankings — yet a substantial proportion of the country's restaurants and bars still operate without formal quality management systems (Rouillon Camino, 2019). Empirical work in the Lima metropolitan area has documented a strong positive correlation between perceived service quality and customer satisfaction ($\rho = 0.866$, $p < .001$) (Ramírez Sulca and Topalaya, 2023), confirming that the precise measurement of this construct carries operational weight, not only academic interest. The Sustainable Development Goals — Goal 8 (decent work) and Goal 12 (responsible consumption) in particular — frame the improvement of service quality in small and medium enterprises as both an economic and a social imperative (UN, 2015). The challenge intensifies in micro-districts where tourism demand, heritage identity and a heterogeneous mix of food-service formats converge — conditions characteristic of the Barranco district in Lima.

At the local level, Barranco represented a highly competitive gastronomic micro-ecosystem in which differentiation depended more on perceived quality and operational consistency than on the product offered. The absence of a structured continuous improvement model and performance indicators generated a visible quality gap between customer expectations and the reality of service delivery. This gap was amplified by the lack of technical training, improvisation in service delivery, and the non-existence of documented processes, all of which affected the long-term sustainability of these businesses. Four structurally distinct establishment types operate within this compact district: fine-dining restaurants, themed restaurants, specialty bars and fast-food outlets. They compete for overlapping visitor and local-resident segments, yet their staffing models, operational complexity and service protocols diverge

sharply. Whether these structural differences produce measurable disparities in the quality of service — as evaluated by the staff who deliver it — had not been examined prior to the present study.

Three gaps in the existing literature motivated this investigation. First, the overwhelming majority of SERVQUAL-based studies adopt the customer perspective, measuring Gap 5 of the original framework — the discrepancy between what customers expect and what they perceive they receive (Parasuraman et al., 1985). Far fewer studies examine Gap 1, the discrepancy between what management believes it delivers and what customers actually expect (Parasuraman et al., 1988). We located no study in the food-and-beverage domain that explicitly operationalises this provider-side assessment across multiple SERVQUAL dimensions. Second, SERVQUAL research in the hospitality sector is dominated by hotels and, to a lesser extent, healthcare settings (Noor, 2025; Arli et al., 2024). Gastronomy — and nightlife-oriented gastronomic economies in particular — remains under-represented, despite evidence that service quality models require adaptation to the specific context in which they are deployed (Sharka et al., 2024). Third, prior empirical work has typically compared service quality across geographic regions or between service categories; no study has compared provider-perceived quality across establishment types within a single gastronomic district, where shared market conditions isolate the effect of organisational structure on quality outcomes.

Against this background, the present study addresses the following research question: how does the service-providing staff of restaurants and bars in the Barranco district perceive the quality of the service they deliver, and which SERVQUAL dimensions present the largest deficits according to establishment type? Three convergent lessons from this review informed the study design: that continuous improvement constitutes a determining factor in achieving sustainable quality standards; that process management and staff training bear directly on customer satisfaction; and that strategic models anchored in performance indicators enable quantifiable gains in productivity and retention. These considerations justified the design of a Strategic Continuous Improvement Model for Service Quality adapted to restaurants and bars in Barranco. Accordingly, the study pursues a two-fold objective: first, to diagnose perceived service quality from the provider perspective across five SERVQUAL dimensions (reliability, responsiveness, assurance, empathy and tangibles) and to identify statistically significant differences among four establishment types; and second, to translate the diagnostic findings into the design of a PDCA-based strategic improvement model (MOO-FS) responsive to the specific quality profile of each establishment category. Figure 1 depicts the conceptual structure of the study. The remainder of the article proceeds with a review of the theoretical foundations, the methodology, results, discussion and conclusions.

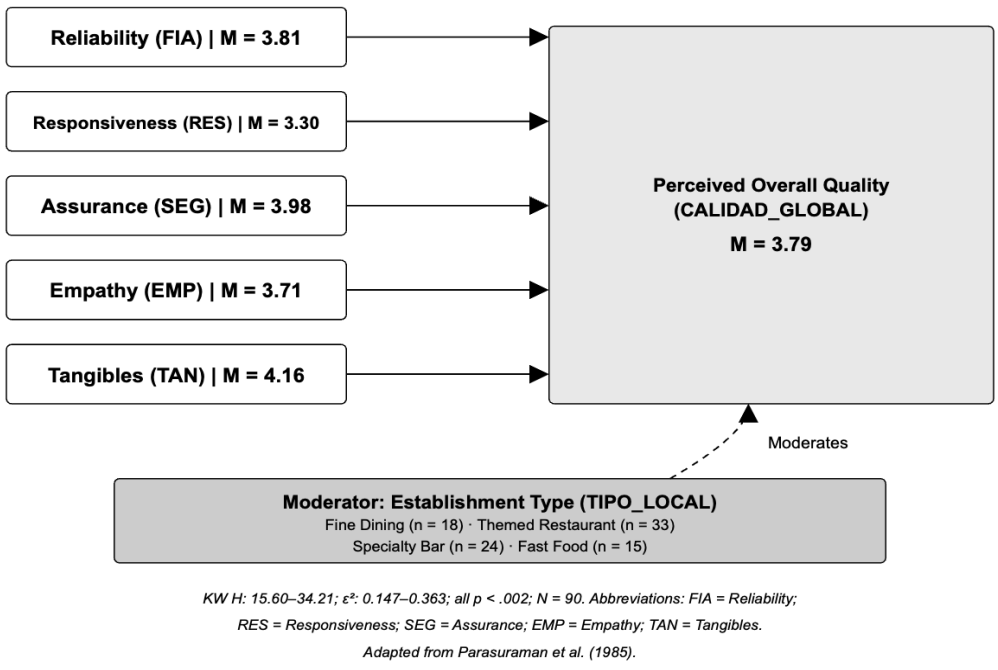

Figure 1. Conceptual structure of the study: adapted SERVQUAL dimensions, overall perceived quality, and the moderating role of establishment type

2. Methodology

2.1 Study design

This study adopted a quantitative, non-experimental, cross-sectional design with a descriptive-comparative scope. It is classified as applied research (Hernández-Sampieri and Mendoza, 2018) and is positioned within Gap 1 of the SERVQUAL framework (Parasuraman et al., 1985), which addresses the discrepancy between what service providers believe they deliver and what customers actually expect. This framing transforms the use of internal staff as informants from a methodological limitation into a theoretically grounded design choice. All propositions are descriptive-comparative; no causal claims are advanced.

2.2 Participants and sampling

The target population comprised food service establishments in the Barranco district of Lima, Peru, a heritage-listed bohemian tourism micro-district. Thirty establishments were selected through non-probability purposive sampling, applying the following inclusion criteria: more than twelve months of continuous operation, minimum seating capacity of twenty, at least five employees, and operation six or more days per week. Within each establishment, three key informants were surveyed — one owner or general manager, one operations manager, and one front-line service employee — yielding a total sample of N = 90. The sample comprised four establishment types: themed restaurants (n = 33, 36.67%), specialty bars (n = 24, 26.67%), fine dining (n = 18, 20.00%) and fast food (n = 15, 16.67%). Data were collected between 30 September and 16 December 2025. Figure 2 presents the participant flow diagram.

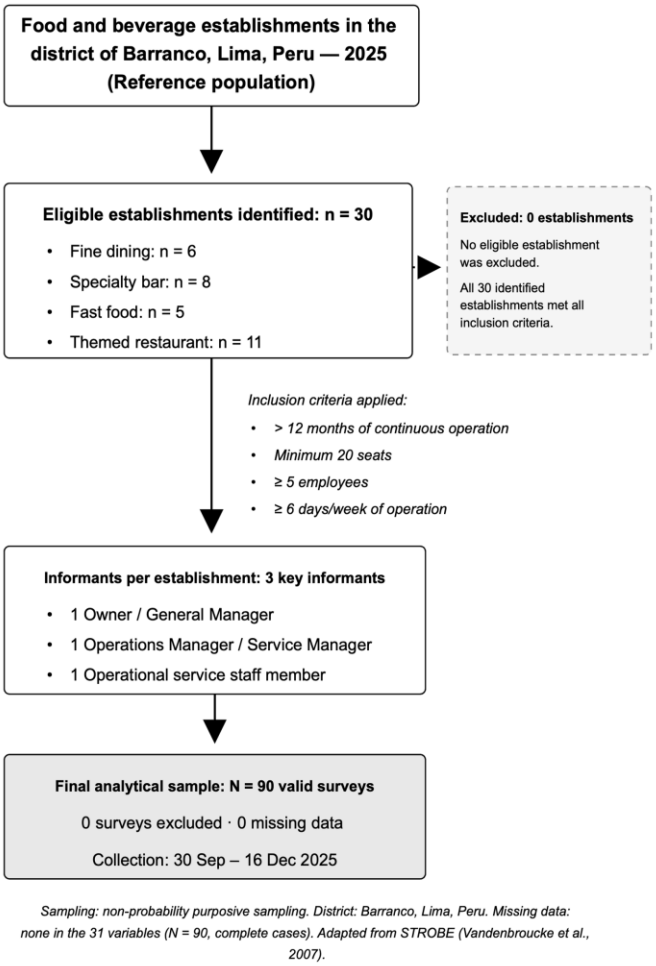

Figure 2. Participant flow diagram (adapted from STROBE guidelines)

2.3 Instrument

The instrument employed was the SERVQUAL questionnaire adapted to the gastronomic sector of Barranco (Parasuraman et al., 1988), structured around five dimensions: reliability (service fulfilment and accuracy), responsiveness (speed and willingness of staff), assurance (trust, knowledge, and professionalism), empathy (personalised treatment and understanding of needs), and tangibles (ambiance, cleanliness, and staff presentation). Each dimension was assessed through two items, yielding a total of 10 closed-ended questions rated on a five-point Likert scale (1 = Strongly disagree, 5 = Strongly agree).

2.4 Pilot test and instrument validation

The pilot test was conducted to validate the clarity, coherence, and reliability of the measurement instrument. For this purpose, 20% of the target population was considered, equivalent to six food service establishments — three restaurants and three bars — located within the Barranco district. Selection followed a purposive sampling approach, prioritising businesses with formal organisational structures and over one year of continuous operation. Participants included owners, operations managers, and service supervisors, as they were the actors directly involved in service quality management.

The instrument was administered in person during the second week of September 2025, with prior formal authorisation from each establishment. The average completion time per survey was 10 minutes, under controlled conditions to minimise response bias. Data were initially recorded and processed in Microsoft Excel 2021, which allowed for review of frequency distributions, item comprehension, and response coherence.

In the reliability analysis phase, Cronbach's alpha was calculated to determine the instrument's internal consistency. The obtained result was $\alpha = 0.87$, a value considered high and acceptable according to the criteria established by George and Mallery (2003), who indicate that an alpha above 0.80 denotes good reliability. This result confirmed that the questions were understood homogeneously and that the questionnaire dimensions exhibited statistical stability. During content validation, expert judges (two research faculty members and one specialist in service quality management) evaluated the items against three criteria: clarity, relevance, and coherence. Results showed an average agreement level of 92%, supporting the instrument's pertinence and its adequacy to the conditions of the Barranco gastronomic environment.

The pilot test led to minor wording adjustments in two items related to the empathy dimension, to improve respondent comprehension. The remaining items retained their original structure, having demonstrated uniform interpretation and high content validity. Accordingly, the adapted SERVQUAL questionnaire was deemed valid and reliable for definitive application across the 30 food service establishments in the district. In the definitive study, the instrument returned a Cronbach's alpha of 0.836 (95% CI [0.781, 0.882], Feldt method), confirming good internal consistency (Doval et al., 2023).

2.5 Data analysis

All analyses were conducted in Python 3.x (SciPy, pingouin, factor_analyzer, statsmodels). The database contained no missing values across all 90 observations and 31 variables; a complete-case strategy was applied throughout. Descriptive statistics (mean, standard deviation, median, range) were computed for each SERVQUAL dimension and for overall perceived quality, both globally and by establishment type. A favourability index classified dimension means ≥ 4.00 as favourable.

Distributional assumptions were assessed via Anderson–Darling and Shapiro–Wilk tests. Five of six variables rejected normality ($p < .005$); only overall perceived quality met the criterion (Shapiro–Wilk $W = 0.978$, $p = .139$). The Kruskal–Wallis H test was therefore adopted as the primary comparison procedure, given its suitability for ordinal, non-normal data (Lukose and Agbeyangi, 2025). Six tests were conducted (one per dimension plus overall quality, $df = 3$). Pairwise post-hoc comparisons used Dunn's test with Bonferroni correction within each SERVQUAL variable (six pairwise contrasts per variable). Effect sizes were quantified as epsilon-squared (ε^2; Cohen, 1988: small 0.01, medium 0.06, large ≥ 0.14), and Cohen's d was computed for each significant pair. Bootstrap BCa 95% confidence intervals for group means were estimated from 10,000 resamples.

The clustering structure (three informants per establishment) was quantified through ICC(1) = 0.586, yielding an effective sample size of $N\_eff \approx 41$. Post-hoc power was estimated at both $N = 90$ and N_eff. The factorial structure of the 10-item instrument was examined through exploratory factor analysis (EFA) with principal axis extraction and oblimin rotation (respondent-to-item ratio 9:1); sampling adequacy was confirmed by KMO = 0.834 and Bartlett's

test ($p < .001$) (Alfatafta et al., 2025; Sharka et al., 2024). For overall perceived quality, a Welch ANOVA served as parametric complement. Two sensitivity checks assessed robustness: (a) Kruskal–Wallis replication on data winsorised at the 5th percentile, and (b) aggregation to the establishment level (N = 30) before re-estimation (Perdomo-Verdecia et al., 2024). Favourability proportions were compared across types using chi-squared or Fisher's exact tests.

2.6 Ethical considerations

The study was conducted in accordance with the Research Ethics Code of the Universidad César Vallejo (2022) and Peruvian data protection legislation (Law 29733). Informed consent was obtained from all participants prior to administration. Participation was voluntary and confidential; no personally identifiable information was recorded.

3. Results

3.1 Sample composition and descriptive statistics

The following sections report the outcomes of that protocol across the 90 observations collected from September to December 2025. The database comprised 90 valid observations with no missing values across all 31 variables (0% missingness; complete-case strategy applied throughout). Themed restaurants contributed the largest share of the sample (n = 33, 36.67%), followed by specialty bars (n = 24, 26.67%), fine dining (n = 18, 20.00%) and fast food (n = 15, 16.67%). Figure 3 displays this distribution.

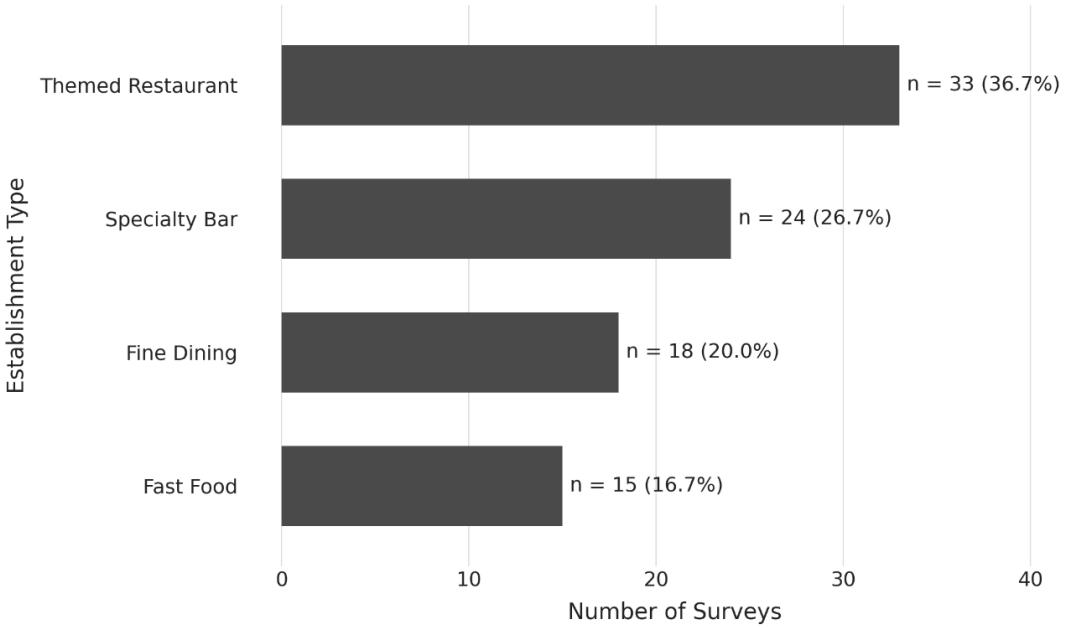

Figure 3. Distribution of Surveys by Establishment Type (N = 90)

The dimension with the highest mean score was tangibles (M = 4.16, SD = 0.61, Mdn = 4.00), reflecting a favourable assessment of ambiance, cleanliness, and staff presentation across the establishments evaluated. Assurance ranked second (M = 3.98, SD = 0.64), followed by reliability (M = 3.81, SD = 0.62) and empathy (M = 3.71, SD = 0.57). The dimension with the lowest mean was responsiveness (M = 3.30, Mdn = 3.00, SD = 0.58). Overall perceived quality yielded a mean of 3.79 (SD = 0.45). The deficit from the scale maximum ($5 - M$) was largest for responsiveness (1.70 points) and smallest for tangibles (0.84 points), establishing a clear intervention hierarchy. Table 1 presents the full descriptive statistics by dimension and establishment type; Figure 4 shows the dimensional profile with bootstrap 95% confidence intervals.

Table 1

Descriptive Statistics for SERVQUAL Dimensions by Establishment Type (N = 90)

	Global (N = 90)			Fine dining (n = 18)		Specialty bars (n = 24)		Fast food (n = 15)		Themed restaurants (n = 33)	
Dimension	**M**	**SD**	**Mdn**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Reliability	3.81	0.62	4.00	4.31	0.55	3.60	0.42	3.70	0.53	3.73	0.70
Responsiveness	3.30	0.58	3.00	3.64	0.51	2.96	0.51	3.70	0.46	3.18	0.54
Assurance	3.98	0.64	4.00	4.58	0.43	3.81	0.62	3.77	0.62	3.86	0.56
Empathy	3.71	0.57	3.50	4.28	0.46	3.52	0.43	3.27	0.42	3.74	0.52
Tangibles	4.16	0.61	4.00	4.78	0.39	3.81	0.44	3.93	0.46	4.18	0.64
Overall quality	3.79	0.45	3.80	4.32	0.26	3.54	0.32	3.67	0.33	3.74	0.44

Note. M = mean; SD = standard deviation; Mdn = median. Means are dimension-level averages of two items each, except Overall quality (average of all five dimensions). BCa 95% confidence intervals computed from 10,000 bootstrap resamples. Groups: Fine dining (n = 18), Specialty bars (n = 24), Fast food (n = 15), Themed restaurants (n = 33). Scores range from 1 to 5.

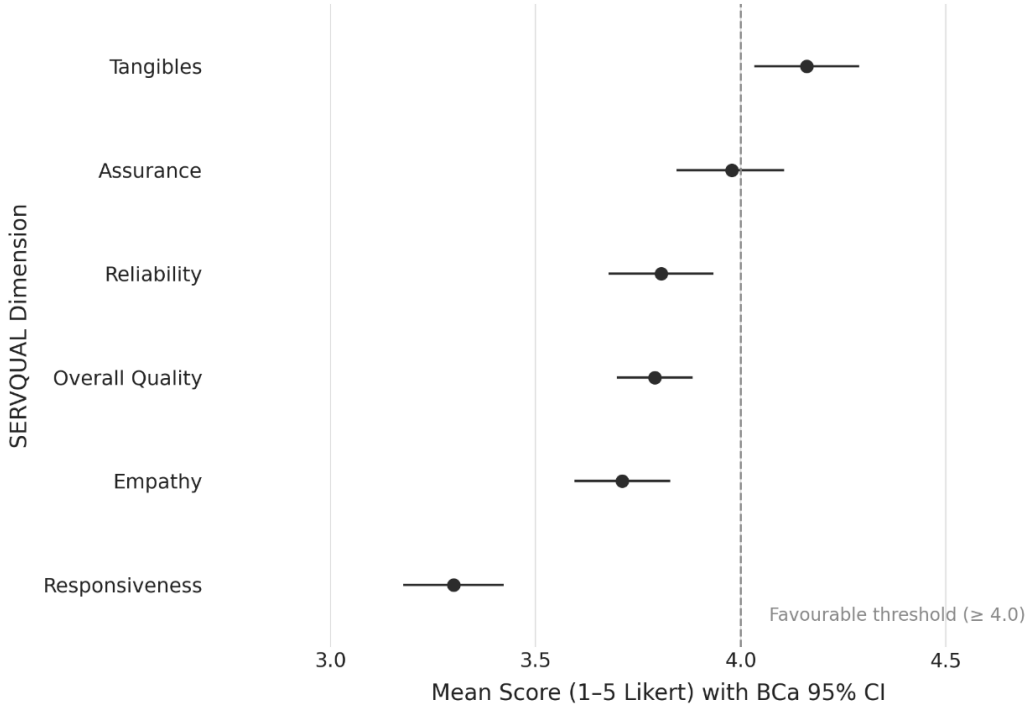

Figure 4. SERVQUAL Dimensional Profile with Bootstrap BCa 95% CIs

When disaggregated by establishment type, fine dining recorded the highest scores on four of five dimensions and overall quality; fast food was numerically highest on responsiveness (M = 3.70 vs. 3.64) (M_global = 4.32, 95% BCa CI [4.18, 4.42]). Specialty bars obtained the lowest overall mean (M = 3.54, 95% BCa CI [3.42, 3.67]). Within specific dimensions, the widest between-group range appeared in tangibles, where fine dining averaged 4.78 against 3.81 for specialty bars — a gap of nearly one full Likert point. Responsiveness showed a distinctive pattern: fast food (M = 3.70) outperformed both specialty bars (M = 2.96) and themed restaurants (M = 3.18), placing it first numerically on responsiveness (M = 3.64), with no significant fine dining–fast food contrast on this dimension.

The favourability analysis classified dimension means ≥ 4.00 as favourable. The dimension with the highest percentage of favourable ratings was tangibles (72.22%), followed by assurance (64.44%) and reliability (53.33%). In contrast, empathy reached 47.78% favourability, while overall perceived quality registered 35.56%. The dimension with the lowest result was responsiveness, with only 26.67% favourable ratings (Figure 5). Chi-squared and Fisher's exact tests confirmed that favourability proportions differed significantly by establishment type for all six variables (all $p < .033$), with Cramér's V ranging from 0.31 (tangibles) to 0.54 (overall quality), indicating medium to large association strength.

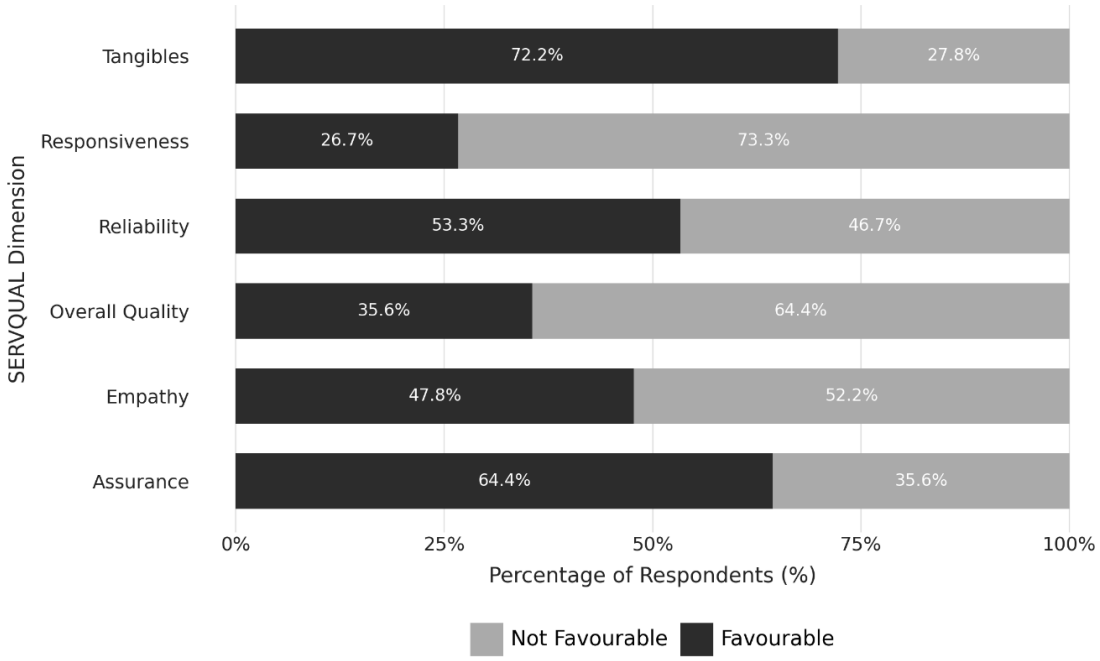

Figure 5. Favourability Index by SERVQUAL Dimension (N = 90)

3.2 Normality diagnostics

Anderson–Darling and Shapiro–Wilk tests indicated that five of six outcome variables departed significantly from normality (all $p < .005$). Reliability ($A^2 = 2.515$, $W = 0.935$, $p < .001$), responsiveness ($A^2 = 3.288$, $W = 0.921$, $p < .001$), assurance ($A^2 = 2.391$, $W = 0.932$, $p < .001$), empathy ($A^2 = 3.000$, $W = 0.927$, $p < .001$) and tangibles ($A^2 = 3.114$, $W = 0.902$, $p < .001$) all exceeded the critical values. Only overall perceived quality met the normality criterion ($A^2 = 0.458$; Shapiro–Wilk $W = 0.978$, $p = .139$). On this basis, the Kruskal–Wallis H test served as the primary comparison procedure for all six variables, with a Welch ANOVA conducted as a parametric complement for overall quality alone. Table 2 presents the complete normality results.

Table 2

Normality Tests and Distributional Assumption Summary (N = 90)

	Anderson–Darling			Shapiro–Wilk		
Variable	**A²**	**Crit. 5%**	**Decision**	**W**	**p**	**Decision**
Reliability	2.515	0.746	Reject	0.935	< .001	Reject
Responsiveness	3.288	0.746	Reject	0.921	< .001	Reject
Assurance	2.391	0.746	Reject	0.932	< .001	Reject
Empathy	3.000	0.746	Reject	0.927	< .001	Reject
Tangibles	3.114	0.746	Reject	0.902	< .001	Reject
Overall quality	0.458	0.746	Do not reject	0.978	.139	Do not reject

Note. AD = Anderson–Darling; W = Shapiro–Wilk statistic. Critical value for Anderson–Darling at 5% = 0.746; at 1% = 1.026. 'Reject' indicates the normality hypothesis was rejected ($p < .05$). 'Do not reject' indicates failure to reject ($p \geq .05$). ICC(1) = 0.586 for overall quality, yielding N_eff ≈ 41. On the basis of these results, Kruskal–Wallis H was used as the primary inferential procedure for all six variables.

3.3 Primary inferential analysis

All six Kruskal–Wallis tests yielded statistically significant differences among establishment types (Table 3). Overall perceived quality showed the strongest association with establishment type ($H(3) = 34.21$, $p < .001$, $\varepsilon^2 = 0.363$), followed by empathy ($H(3) = 29.97$, $p < .001$, $\varepsilon^2 = 0.314$), tangibles ($H(3) = 29.07$, $p < .001$, $\varepsilon^2 = 0.303$), responsiveness ($H(3) = 22.84$, $p < .001$, $\varepsilon^2 = 0.231$), assurance ($H(3) = 20.77$, $p < .001$, $\varepsilon^2 = 0.207$) and reliability ($H(3) = 15.60$, $p = .001$, $\varepsilon^2 = 0.147$). Figure 6 presents the group means with BCa 95% confidence intervals for all six variables.

Table 3

Kruskal–Wallis H Tests Comparing SERVQUAL Dimensions Across Establishment Types (N = 90, df = 3)

Variable	**N**	**df**	**H**	**p**	**ε²**	**Effect magnitude**
Overall quality	90	3	34.21	< .001	0.363	Large
Empathy	90	3	29.97	< .001	0.314	Large
Tangibles	90	3	29.07	< .001	0.303	Large
Responsiveness	90	3	22.84	< .001	0.231	Large
Assurance	90	3	20.77	< .001	0.207	Large
Reliability	90	3	15.60	.001	0.147	Large

Note. ε^2 = epsilon-squared (effect size for Kruskal–Wallis). Effect magnitude thresholds: small $\varepsilon^2 \geq .01$; medium $\varepsilon^2 \geq .06$; large $\varepsilon^2 \geq .14$ (Tomczak and Tomczak, 2014). All p values two-tailed. Complete-case analysis; N excluded = 0 for all variables. Results ordered by ε^2 descending.

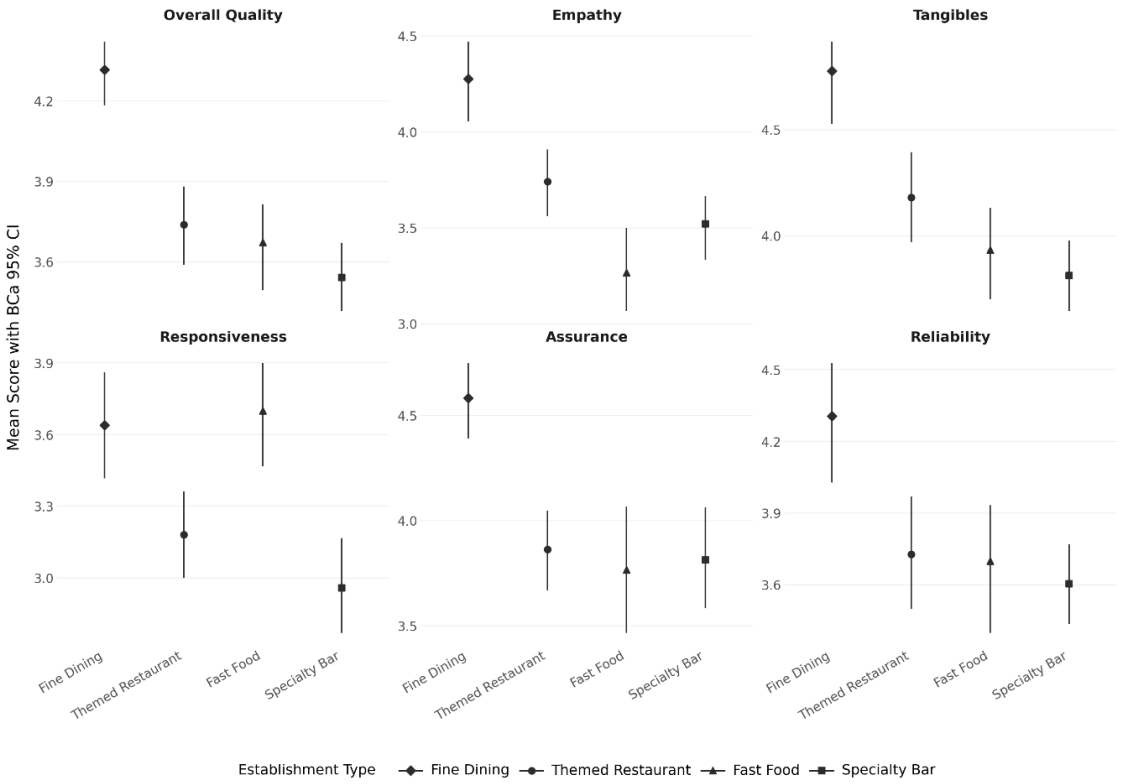

Figure 6. Group Means and BCa 95% CIs by Establishment Type (Six SERVQUAL Variables)

Every effect size fell in the large range ($\varepsilon^2 \geq 0.14$). Figure 7 displays this effect-size hierarchy, confirming overall quality and empathy as the two variables most strongly differentiated by establishment type, and reliability as the comparatively smallest, though still large, effect.

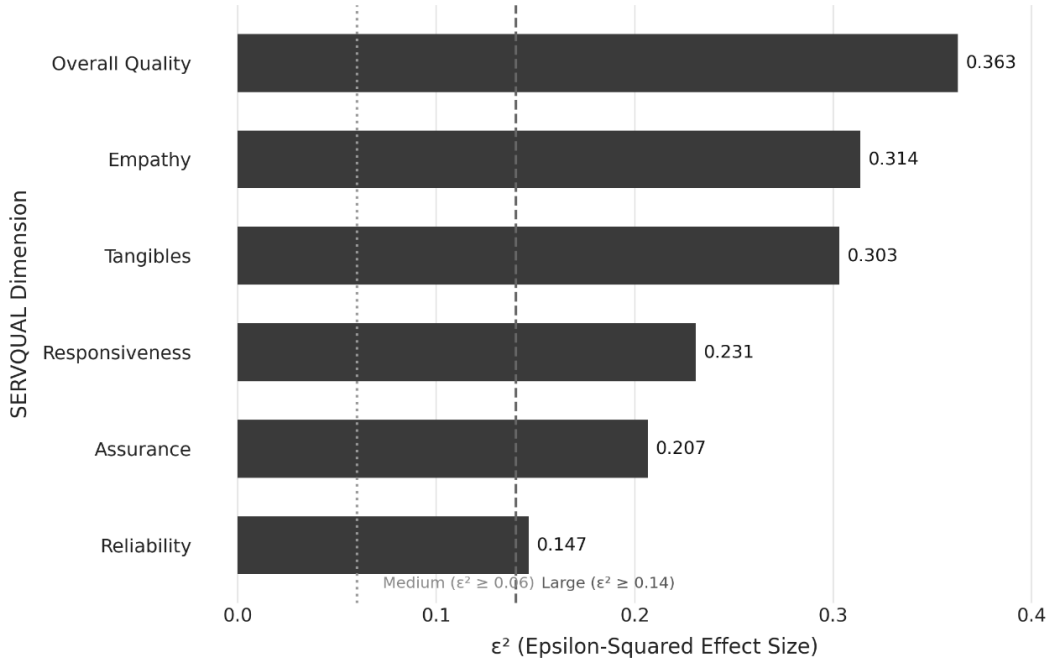

Figure 7. Effect Sizes (Epsilon-Squared) for Kruskal–Wallis Tests by SERVQUAL Dimension

Dunn–Bonferroni post-hoc comparisons (Table 4) identified 18 statistically significant contrasts out of 36 tested. Fine dining differed significantly from all other types on assurance, empathy, tangibles and overall quality; on reliability it differed from specialty bars and themed restaurants but not from fast food; on responsiveness it differed only from specialty bars, with Cohen's d values consistently in the large range (d = 0.89 to 2.60). The largest pairwise effect was observed for overall quality between fine dining and specialty bars (d = 2.60, p_adj < .001). On responsiveness, specialty bars recorded the lowest rank mean and differed significantly from fast food (d = −1.52, p_adj < .001); fast food also outperformed themed restaurants on this dimension (d = 1.00, p_adj = .027). For empathy, fine dining outperformed all three groups, with the strongest contrast against fast food (d = 2.29, p_adj < .001); fast food in turn scored significantly below themed restaurants (d = −0.97, p_adj = .026). No significant differences emerged between specialty bars and themed restaurants on any dimension after Bonferroni correction, nor between specialty bars and fast food on reliability, assurance or tangibles.

Table 4

Significant Dunn–Bonferroni Pairwise Comparisons (p_adj < .05; 18 of 36 Pairs Tested)

Variable	Group 1	Group 2	z	p_adj	d	Magnitude
Reliability	Fine dining	Specialty bars	3.726	.001	1.474	Large
Reliability	Fine dining	Themed restaurants	3.240	.007	0.892	Large
Responsiveness	Fine dining	Specialty bars	3.739	.001	1.337	Large
Responsiveness	Specialty bars	Fast food	-3.942	< .001	-1.516	Large
Responsiveness	Fast food	Themed restaurants	2.836	.027	1.002	Large
Assurance	Fine dining	Specialty bars	3.961	< .001	1.406	Large
Assurance	Fine dining	Fast food	3.624	.002	1.554	Large
Assurance	Fine dining	Themed restaurants	3.856	< .001	1.383	Large
Empathy	Fine dining	Specialty bars	4.189	< .001	1.709	Large
Empathy	Fine dining	Fast food	5.132	< .001	2.290	Large
Empathy	Fine dining	Themed restaurants	3.086	.012	1.074	Large
Empathy	Fast food	Themed restaurants	-2.858	.026	-0.973	Large
Tangibles	Fine dining	Specialty bars	5.139	< .001	2.305	Large
Tangibles	Fine dining	Fast food	3.997	< .001	1.997	Large

Tangibles	Fine dining	Themed restaurants	3.323	.005	1.059	Large
Overall quality	Fine dining	Specialty bars	5.649	< .001	2.597	Large
Overall quality	Fine dining	Fast food	4.079	< .001	2.190	Large
Overall quality	Fine dining	Themed restaurants	4.259	< .001	1.501	Large

Note. d = Cohen's d (effect size for pairwise difference computed from raw group means using pooled standard deviations). p_adj = Bonferroni-adjusted p value. Effect magnitude: small |d| ≥ 0.20; medium |d| ≥ 0.50; large |d| ≥ 0.80 (Cohen, 1988). Negative z and d indicate Group 2 rank mean > Group 1. All 36 pairs tested; 18 remained significant after Bonferroni correction. Non-significant pairs (n = 18) omitted from this table. Groups: Fine dining (n = 18); Specialty bars (n = 24); Fast food (n = 15); Themed restaurants (n = 33).

3.4 Factorial structure of the instrument

Exploratory factor analysis with oblimin rotation extracted five factors that jointly accounted for 48.5% of the total item variance (Table 5). The Kaiser–Meyer–Olkin index (0.834) and Bartlett's test ($p < .001$) confirmed adequate sampling. Factor 1 loaded primarily on Q10_TAN (0.72) and Q9_TAN (0.59), corresponding to the tangibles dimension. Factor 2 loaded on Q3_RES (0.92) and, to a lesser degree, Q4_RES (0.42), mapping onto responsiveness. Factor 3 captured reliability through Q2_FIA (0.92) and Q1_FIA (0.46). Factor 5 loaded on Q6_SEG (0.81), the assurance dimension. Factor 4 showed weaker and dispersed loadings. The structure partially mirrors the five theoretical SERVQUAL dimensions, with tangibles, responsiveness and reliability forming the most clearly differentiated factors.

Table 5

EFA Oblimin-Rotated Factor Loadings (N = 90, 10 Items, 5 Factors)

Item	**Dimension**	**F1**	**F2**	**F3**	**F4**	**F5**
Q1_FIA	*Reliability*	-0.060	0.101	**0.462**	-0.235	0.113
Q2_FIA	*Reliability*	0.033	-0.005	**0.925**	0.026	0.017
Q3_RES	*Responsiveness*	-0.027	**0.916**	0.019	0.018	-0.010
Q4_RES	*Responsiveness*	0.188	**0.418**	-0.098	-0.190	-0.029
Q5_SEG	*Assurance*	0.298	0.028	0.061	-0.381	0.282
Q6_SEG	*Assurance*	-0.004	0.015	0.076	-0.005	**0.805**
Q7_EMP	*Empathy*	0.150	0.395	-0.018	-0.008	0.300
Q8_EMP	*Empathy*	**0.477**	0.112	0.008	0.322	0.159
Q9_TAN	*Tangibles*	**0.591**	-0.042	-0.044	-0.085	0.240

Q10_TAN	*Tangibles*	**0.723**	0.068	0.172	-0.035	-0.094
SS loadings		1.252	1.201	1.121	0.350	0.924
Prop. variance		0.125	0.120	0.112	0.035	0.092
Cumulative variance		0.125	0.245	0.357	0.392	0.485

Note. Extraction method: principal axis factoring. Rotation: oblimin (oblique). Loadings ≥ |.40| are shown in bold and considered salient. F1 = Tangibles; F2 = Responsiveness; F3 = Reliability; F4 = Mixed/weak; F5 = Assurance. KMO overall = 0.834; Bartlett's test of sphericity $p < .001$. Cumulative variance explained = 48.5%. SS loadings = sum of squared loadings per factor.

3.5 Sensitivity and robustness checks

Three robustness checks corroborated the primary findings (Figure 8). First, Kruskal–Wallis replication on data winsorised at the 5th percentile preserved statistical significance for every dimension; H and ε^2 values changed by less than 2% from the original estimates (e.g., responsiveness: $H_{winsorised} = 23.17$ vs. $H_{original} = 22.84$, $\Delta\varepsilon^2 = 1.7\%$). Second, aggregation to the establishment level (N = 30, one mean per establishment) returned significant results for all six variables (all $p < .021$), with effect sizes increasing in magnitude (ε^2 range: 0.261–0.525), as expected when within-establishment variance is absorbed. Third, a Welch ANOVA for overall perceived quality confirmed the non-parametric result ($F(3, 42.74) = 27.41$, $p < .001$); Games–Howell post-hoc tests replicated the Dunn pattern, with fine dining differing significantly from all other types (all $p < .001$) and no significant contrasts among the remaining three.

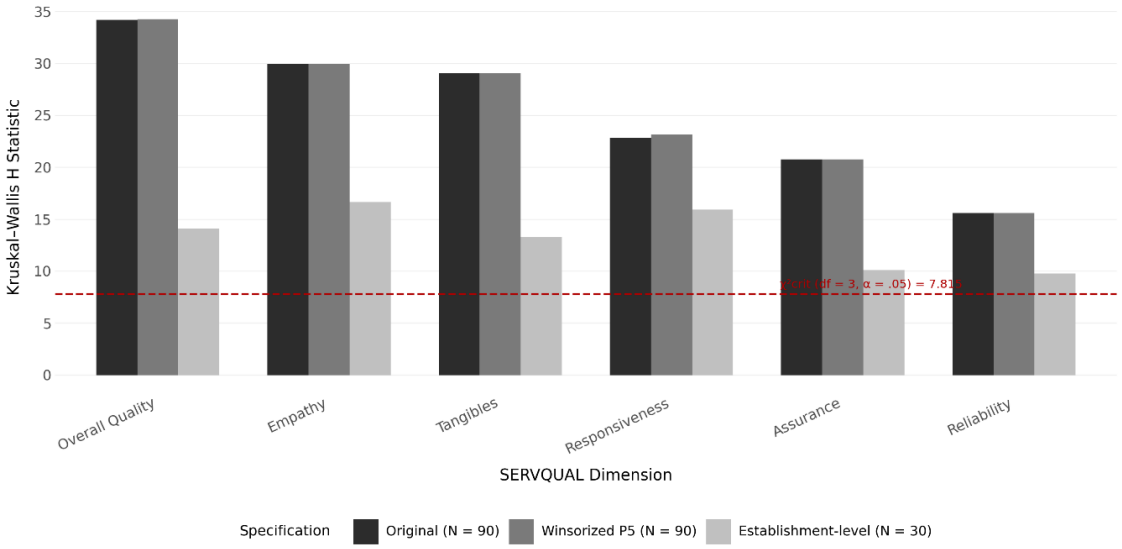

Figure 8. Sensitivity Analysis: Kruskal–Wallis H Statistic Across Three Specifications

Post-hoc power analysis indicated that at the nominal sample (N = 90) power exceeded 0.91 for all six tests. At the effective sample size ($N_{eff} \approx 41$, adjusted for clustering via ICC(1) = 0.586), power exceeded 0.75 for four of the five remaining dimensions; assurance reached 0.75 and reliability ($\varepsilon^2 = 0.147$, power = 0.55 at N_{eff}) fell below the conventional 0.80 threshold, consistent with its position as the smallest observed effect. The minimum detectable effect at N_{eff} was $\varepsilon^2 = 0.226$, confirming adequate sensitivity for the medium-to-large effects documented across the remaining variables.

4. Discussion

The present study set out to diagnose how the service-providing staff of restaurants and bars in the Barranco district perceive the quality they deliver, and to determine which SERVQUAL dimensions present the largest deficits by establishment type. Four principal findings warrant discussion: responsiveness emerged as the weakest dimension; tangibles as the strongest; all six Kruskal–Wallis tests returned large effects differentiating the four establishment categories; and overall perceived quality fell below the favourable threshold. We examine each finding in turn, contrasting it with the existing literature and acknowledging the boundary conditions that frame its interpretation.

Responsiveness recorded the lowest mean among all dimensions ($M = 3.30$; $\varepsilon^2 = 0.231$) and the lowest favourability rate (26.67%). In healthcare settings, Rahmatia et al. (2025) found that empathy, not responsiveness, carried the strongest effect on patient satisfaction ($\beta = 0.629$, $p < .001$); in dental services, Sharka et al. (2024) reported a more modest responsiveness coefficient ($\beta = 0.181$). Our result aligns more closely with Sharka's pattern, though the context differs substantially: in a nightlife-oriented gastronomic district, peak-hour demand surges and limited kitchen-to-table coordination may compress service times beyond what staffing levels can absorb. That the providers themselves identify speed and willingness as their primary deficit is telling — it suggests that Gap 1 awareness exists at the operational level but is not being translated into managerial action (Parasuraman et al., 1985). A distinctive sub-finding reinforces this reading: on responsiveness, fast food ($M = 3.70$) significantly outperformed specialty bars ($M = 2.96$, $d = -1.52$) and themed restaurants ($M = 3.18$, $d = 1.00$). Fast-food operations are structurally designed for speed — standardised workflows, limited menus, defined throughput targets — which may explain why their staff rate responsiveness more favourably than staff in formats where service pacing is less predictable. A question remains about whether customers would rank responsiveness similarly; the provider-only lens of this study, framed within Gap 1, captures one side of that equation.

Tangibles, by contrast, recorded the highest global mean ($M = 4.16$; 72.22% favourable) and the smallest deficit from the scale maximum (0.84 points). This finding converges with Alfatafta et al. (2025), who reported tangibility as the highest-scoring dimension ($M = 4.10$) in Palestinian rehabilitation services, and with Bhobet and Baluyot (2025), who documented favourable tangible quality in IoT-enabled hospitality settings. In the Barranco context, physical infrastructure — ambiance, cleanliness, staff presentation — appears to be the dimension where providers' self-assessment most closely approximates what a reasonable customer might expect: the narrowest Gap 1. Fine dining registered a median of 5.00 on tangibles, a ceiling effect that signals full provider confidence in physical standards. From a resource-allocation perspective, Perdomo-Verdecia et al. (2024) cautioned that equal investment across all service attributes is operationally inefficient; our deficit hierarchy supports this principle by indicating that redirecting resources from an already strong tangibles dimension toward responsiveness and empathy would likely yield higher marginal returns.

The third finding — that establishment type functions as a strong differentiator of perceived quality across every dimension ($\varepsilon^2 = 0.147$–0.363, all large) — is perhaps the most consequential for policy. It challenges the assumption, common in district-level hospitality management, that a uniform quality protocol can serve structurally heterogeneous food-service businesses. Perdomo-Verdecia et al. (2024) reached a parallel conclusion in a hotel context, showing that equifinal configurations rather than a single formula produce satisfaction. We extend this logic to gastronomy: within a single geographic cluster, shared market conditions do not homogenise quality profiles. Fine dining outperformed all other types on most variables ($d = 0.89$–2.60), while specialty bars and themed restaurants occupied statistically indistinguishable middle positions. These patterns held across three analytic specifications — original, winsorised, and establishment-level — which is reassuring given the clustered sampling structure ($ICC(1) = 0.586$, $N_{eff} \approx 41$). The effective sample size does constrain power for the smallest observed effect (reliability, $\varepsilon^2 = 0.147$, power = 0.55 at N_{eff}); multilevel modelling, following Al-refaei et al. (2023), would formally partition within- and between-establishment variance and represents the natural analytic extension.

Beneath these group-level differences lies a question of measurement: does the 10-item instrument actually reproduce the five-factor structure that SERVQUAL postulates? The exploratory factor analysis addresses this directly. The five extracted factors accounted for 48.5% of item variance — below the 60% threshold sometimes applied in confirmatory contexts but within the range that Bulhões et al. (2023) described as acceptable in exploratory behavioural research. Tangibles, responsiveness and reliability formed the most clearly differentiated factors, while empathy and assurance items showed cross-loadings, a pattern consistent with the broader SERVQUAL literature where these two dimensions frequently merge in non-Western settings (Noor, 2025). Sharka et al. (2024) similarly reported dimensional overlap when adapting SERVQUAL to dental services, noting that context-specific models may need fewer or reconfigured dimensions. In the present case, the reduced 10-item instrument (respondent-to-item ratio

9:1) was designed as a pragmatic field tool; the EFA should be read as exploratory confirmation of adequate factorial behaviour rather than definitive validation. A confirmatory factor analysis with $N \geq 200$ remains the priority extension.

Overall perceived quality registered a mean of 3.79 (SD = 0.45), with only 35.56% of responses crossing the favourable threshold. This composite signal carries theoretical weight: even the staff who deliver the service rate it below the benchmark that conventional customer-facing SERVQUAL studies commonly report (means above 3.80 appear in Olawuyi and Kleynhans, 2025 and Cayo-Velásquez et al., 2025). The provider-side result constitutes diagnostic evidence for the existence of Gap 1 awareness in the Barranco context — staff recognise the shortfall, yet the absence of structured quality management systems means this recognition does not translate into corrective action. The cross-sectional design captures this awareness at a single point in the district's operational cycle (September–December 2025); whether the deficit persists, deepens, or narrows after intervention can only be determined through longitudinal follow-up (Hernández-Sampieri and Mendoza, 2018).

From a practical standpoint, the dimension-specific deficit hierarchy provides the empirical input that a Plan–Do–Check–Act framework requires. Endalamaw et al. (2024) observed that continuous quality improvement succeeds when it targets specific process weaknesses rather than diffuse quality aspirations; Naughton et al. (2024) demonstrated that structured, checklist-driven models are compatible with regulated service environments. Applied to Barranco, these principles indicate that the proposed MOO-FS model should be modular: responsiveness protocols for fast food and specialty bars (where the deficit is most acute), empathy-training cycles for fast food (the sole type where empathy was significantly lower than in themed restaurants, $d = -0.97$), and a 'protect and monitor' stance for tangibles across all segments. Fine dining's quality profile serves as an internal benchmark rather than an intervention target, offering the remaining establishment types a concrete, locally grounded standard to approximate.

5. Conclusion

This study diagnosed provider-perceived service quality in 30 food-and-beverage establishments in the Barranco district of Lima, using an adapted 10-item SERVQUAL instrument administered to 90 internal informants. Responsiveness emerged as the dimension with the largest deficit from the scale maximum (1.70 points; M = 3.30), while tangibles constituted the primary strength (M = 4.16; 72.22% favourable). Overall perceived quality fell below the favourable threshold (M = 3.79; 35.56% favourable). All six Kruskal–Wallis tests confirmed that establishment type is a strong differentiator of perceived quality, with uniformly large effect sizes ($\varepsilon^2 = 0.147$–0.363) and 18 of 36 pairwise comparisons reaching significance after Bonferroni correction. Fine dining showed the strongest overall quality profile and led on four of five dimensions; fast food was numerically highest on responsiveness.

The study contributes to the hospitality literature in three respects. First, it operationalises SERVQUAL from the provider perspective (Gap 1) in a gastronomic context — an approach validated by emerging instruments such as SERVQUAL-E but not previously applied to food-and-beverage nightlife economies. The finding that staff themselves rate overall quality below the favourable threshold constitutes evidence that Gap 1 awareness exists but is not being channelled into structured improvement. Second, the intra-district comparison across four establishment types demonstrates that a single geographic cluster does not produce uniform quality profiles, a result that complicates district-level policy prescriptions and calls for modular, type-specific intervention designs. Third, the dimension-specific deficit hierarchy supplies the empirical foundation for the proposed MOO-FS model, enabling PDCA cycles calibrated to each establishment category rather than imposed uniformly across the district.

Several boundary conditions apply. The cross-sectional, single-district, non-probability design establishes a diagnostic baseline rather than generalisable estimates. The provider-only perspective captures Gap 1 but does not directly measure customer expectations. The reduced instrument, while internally consistent ($\alpha = 0.836$) and factorially adequate (KMO = 0.834), awaits confirmatory validation with a larger sample. The clustered sampling structure (ICC = 0.586, $N_{eff} \approx 41$) constrains effective power for the smallest observed effects. The consistency of results across multiple analytic specifications — parametric and non-parametric, full-sample and winsorised, individual and establishment-level — lends confidence to the core conclusions.

Three avenues for future research emerge as priorities. A dual-perspective study pairing provider and customer assessments would yield a direct operationalisation of Gap 1 in the Barranco gastronomic sector. Confirmatory factor analysis with $N \geq 200$ would strengthen the psychometric basis of the adapted scale and permit formal comparison with the 22-item original. Multilevel modelling would explicitly partition within-establishment and between-establishment variance, moving the analysis from descriptive comparison to structural explanation. Together, these extensions would transform the diagnostic baseline documented here into a monitoring framework capable of evaluating the impact of the MOO-FS model once implemented.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Alfatafta, H., Mrayyan, M.T., Al-Atiyyat, N., Al-Rawashdeh, S., Aljezawi, M., Khatib, A., Aldalaykeh, M., Rababa, M., Alhroub, A. and Fawares, A. (2025) 'Assessing service quality and its impact on patient experience and patient satisfaction using structural equation modelling', BMC Health Services Research, Vol. 25, No. 1, Article 300, DOI: 10.1186/s12913-025-13172-z.
2. Al-refaei, A.A., Zumrah, A.R., Alsamawi, M.S. and Alasali, A.N. (2023) 'An integrated mediating and moderating model to improve service quality through job involvement, job satisfaction, and organisational commitment', Sustainability, Vol. 15, No. 10, Article 7978, DOI: 10.3390/su15107978.
3. Arli, D., Bauer, C. and Palmatier, R.W. (2024) 'The impact of SERVQUAL on consumers' satisfaction, trust, and loyalty: a study on restaurants', Journal of Promotion Management, Vol. 30, No. 8, pp.1275–1302, DOI: 10.1080/10496491.2024.2372858.
4. Bhobet, L.A.R. and Baluyot, M.B. (2025) 'Technological transformation in hospitality: impact of IoT-enabled service environments on perceived quality and customer satisfaction', Journal of Informatics and Web Engineering, Vol. 10, No. 43s, Article 8359, DOI: 10.52783/jisem.v10i43s.8359.
5. Bulhões, L.A., Gasques, A.C.F., Giordani da Silveira, G.A., Pallaro, M.A., Moreira da Silva, F.A. and Borille, A.V. (2023) 'Evaluation of waste in food services: a structural equation model', Sustainability, Vol. 15, No. 10, Article 8044, DOI: 10.3390/su15108044.
6. Cayo-Velásquez, D.R., Solano Salinas, C.J. and Huamán Soto, G.J. (2025) 'Review of the concept of service quality and its measurement models in Latin America', Journal of Ecohumanism, Vol. 4, No. 2, Article 5817, DOI: 10.62754/joe.v4i2.5817.
7. Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences, 2nd ed., Lawrence Erlbaum Associates, Hillsdale, NJ.
8. Doval, E., Viladrich, C. and Angulo-Brunet, A. (2023) 'Coefficient alpha: the resistance of a classic', Psicothema, Vol. 35, No. 1, pp.5–20, DOI: 10.7334/psicothema2022.321.
9. Endalamaw, A., Khatri, R.B., Mengistu, T.S., Erku, D., Wolka, E., Zewdie, A., Assefa, Y. and Gilks, C.F. (2024) 'A scoping review of continuous quality improvement in healthcare system: conceptualisation, models and tools', BMC Health Services Research, Vol. 24, No. 1, Article 487, DOI: 10.1186/s12913-024-10828-0.
10. George, D. and Mallery, P. (2003) SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 Update, 4th ed., Allyn and Bacon, Boston, MA.
11. Hernández-Sampieri, R. and Mendoza, C.P. (2018) Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta, McGraw-Hill, Ciudad de México.
12. Lukose, J.M. and Agbeyangi, A.O. (2025) 'Enhancing ICT literacy and sustainable practices in higher education institutions using a service quality approach', Journal of Infrastructure, Policy and Development, Vol. 9, No. 7, Article 11425, DOI: 10.24294/jipd11425.
13. Naughton, B., Zujewska, S.K. and O'Connor, A. (2024) 'A structured model for continuous improvement methods for the pharmaceutical industry', Heliyon, Vol. 10, No. 21, Article e40034, DOI: 10.1016/j.heliyon.2024.e40034.
14. Noor, A.Y.M. (2025) 'Halal service quality: systematic review, conceptualisation and future research agenda', Journal of Islamic Accounting and Business Research, [online] DOI: 10.1108/jiabr-06-2024-0211.
15. Olawuyi, B.J. and Kleynhans, I.C. (2025) 'A bibliometric analysis of service quality in the hospitality industry', Administrative Sciences, Vol. 15, No. 6, Article 215, DOI: 10.3390/admsci15060215.
16. Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985) 'A conceptual model of service quality and its implications for future research', Journal of Marketing, Vol. 49, No. 4, pp.41–50, DOI: 10.1177/002224298504900403.
17. Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988) 'SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality', Journal of Retailing, Vol. 64, No. 1, pp.12–40.
18. Perdomo-Verdecia, V., Mondaca-Marino, C. and González-Cancelas, N. (2024) 'An fsQCA analysis of service quality for hotel customer satisfaction', International Journal of Hospitality Management, Vol. 120, Article 103793, DOI: 10.1016/j.ijhm.2024.103793.
19. Rahmatia, H., Onsardi, O., Harsono, I. and Dharma, F. (2025) 'Service quality in hospital inpatient care: SERVQUAL-based assessment of patient experience and satisfaction', Health SA Gesondheid, Vol. 30, No. 1, Article 3055, DOI: 10.4102/hsag.v30i0.3055.
20. Ramírez Sulca, Y. and Topalaya Berrocal, L.G. (2023) Calidad de servicio y satisfacción de los clientes en un restaurante, Santa Anita, Lima 2023 [Service quality and customer satisfaction in a restaurant, Santa Anita, Lima 2023], undergraduate thesis, Universidad César Vallejo, Lima [online] https://hdl.handle.net/20.500.12692/156443.
21. Rouillon Camino, M.L. (2019) Diagnóstico del área de cocina de Pymes del sector gastronómico para la gestión de calidad: el caso del restaurante "Tabla Caliente" en el 2018 [Kitchen-area diagnosis of gastronomic SMEs for quality management: the case of the 'Tabla Caliente' restaurant in 2018], undergraduate thesis, Pontificia Universidad Católica del Perú, Lima [online] http://hdl.handle.net/20.500.12404/13503.

22. Sharka, R., Abed, H., Tounsi, A. and Tabbaa, D. (2024) ‘Measuring the impact of dental service quality on patient satisfaction using an adapted SERVQUAL framework’, Frontiers in Oral Health, Vol. 5, Article 1362659, DOI: 10.3389/froh.2024.1362659.
23. Tomczak, M. and Tomczak, E. (2014) ‘The need to report effect size estimates revisited: an overview of some recommended measures of effect size’, Trends in Sport Sciences, Vol. 21, No. 1, pp.19–25.
24. UN (2015) Transforming Our World: The 2030 Agenda for Sustainable Development, United Nations, New York, accessed 8 May 2026, https://sdgs.un.org/.
25. Universidad César Vallejo (2022) Código de Ética en Investigación [Research Ethics Code], Universidad César Vallejo, Lima.