Expectation, Perception, and Confirmation: A Mixed-Methods Study of Academic Library Service Quality and Student Satisfaction in Chongqing, China

Zhang Cheng[1] and Subadrah Madhawa Nair[2]

[1,2]Faculty of Education and Humanities, UNITAR International University, Petaling Jaya, Malaysia

Corresponding author:

DOI : https://doi.org/10.47772/IJRISS.2026.102400006

Received: 19 August 2026; Accepted: 24 August 2026; Published: 25 September 2026

ABSTRACT

Chinese universities have invested heavily in library infrastructure over the past decade, yet student evaluations continue to fall short of institutional expectations. This study examines why, by treating the distance between what students expect of an academic library and what they perceive it to deliver as the analytic object rather than as background noise. Using an explanatory sequential mixed-methods design grounded in the SERVQUAL model and Expectation–Confirmation Theory, questionnaire data were collected from 400 undergraduates at five universities in Chongqing, followed by semi-structured interviews analysed thematically. Paired-samples t-tests revealed statistically significant negative gaps for overall service quality and for each of the five SERVQUAL dimensions, with the overall gap reaching an unusually large magnitude (mean difference = 0.401, $t(399) = 74.58$, $p < .001$). All five perceived dimensions correlated positively with satisfaction, and together they explained 34.5% of its variance, with Tangibility emerging as the strongest predictor ($\beta = .304$), followed by Empathy ($\beta = .269$). The qualitative strand produced five themes that mapped directly onto the SERVQUAL dimensions and four improvement priorities: facilities, resources, staff attitude and training, and digital support. Integration showed that students' expectations were anchored not in their own institution's past performance but in externally sourced comparisons (commercial digital platforms and better-resourced universities), which explains why negative disconfirmation persisted despite documented improvements in service provision. The study contributes a contextualised account of how the relative weight of SERVQUAL dimensions shifts with regional infrastructure conditions, and argues that gap-based diagnosis remains more actionable for library managers than perception-only satisfaction measurement.

Keywords: academic library service quality, SERVQUAL, Expectation–Confirmation Theory, student satisfaction, mixed methods, Chinese higher education

INTRODUCTION

Academic libraries in China occupy an unusual position. They have been the beneficiaries of sustained capital investment, particularly since the expansion of higher education enrolment, and they are simultaneously the object of persistent student criticism. Institutional reporting tends to describe libraries in terms of inputs such as square metres of study space, volumes held, and databases subscribed, while student evaluations register something different, and the two accounts rarely meet. This study is concerned with the space between them.

The dominant instrument for measuring that space remains SERVQUAL (Parasuraman et al., 1988), which conceptualises service quality not as a level of performance but as a discrepancy: the difference between what users expect and what they perceive they receive. The model's diagnostic power lies precisely in this relational definition. A library that scores moderately well on perceived performance may still generate dissatisfaction if its users expected considerably more, and a library scoring less well may satisfy users

whose expectations are modest. Perception-only satisfaction surveys, which remain common in library assessment practice, cannot distinguish between these situations and therefore cannot tell managers where intervention would produce the greatest return.

Two limitations characterise the existing evidence base. First, empirical work on library service quality in China has clustered around the coastal and metropolitan institutions Beijing, Shanghai, Jiangsu, Wuhan (Gyau et al., 2021; Peng et al., 2022) while the rapidly expanding inland education hubs have received comparatively little attention. Chongqing, a municipality with a resident population above thirty million and a university student population exceeding one million, is one such setting. Its libraries operate under a distinctive pressure: student numbers have grown faster than physical capacity, and the resulting congestion shapes evaluation in ways that may not be visible in better-resourced contexts. Second, and more consequentially for theory, a large share of SERVQUAL-based library studies report gap statistics without explaining them. Gaps are documented, dimensions are ranked, and recommendations follow but the question of what students were comparing against when they formed their expectations is left unaddressed. Expectations are treated as given rather than as constructed.

This study addresses both limitations. It applies SERVQUAL alongside Expectation–Confirmation Theory (Oliver, 1980) in an explanatory sequential mixed-methods design, using the qualitative strand not to illustrate the quantitative findings but to interrogate them; specifically, to recover the reference points against which students judged their library. Three questions guide the analysis:

1. Do significant differences exist between students' expectations and their perceptions of library service quality, overall and across the five SERVQUAL dimensions?

2. To what extent do perceptions of Tangibility, Reliability, Responsiveness, Assurance, and Empathy relate to and predict student satisfaction?

3. How do students account for their evaluations of library service quality, and what improvements do they prioritise?

The contribution is threefold. Empirically, the study provides the first multi-institution gap analysis of university libraries in Chongqing. Theoretically, it shows that the relative weight of SERVQUAL dimensions is context-dependent in a specifiable way, and it advances a mechanism externally anchored expectation formation that accounts for persistent negative disconfirmation in a period of rising investment. Methodologically, it demonstrates what gap-based measurement adds to perception-only assessment.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Service Quality as Discrepancy

The SERVQUAL model (Parasuraman et al., 1988) specifies five dimensions along which users evaluate services. Tangibility concerns physical facilities, equipment, and visible resources; Reliability, the accurate and dependable delivery of promised services; Responsiveness, the willingness to provide prompt assistance; Assurance, the competence and courtesy of staff and the confidence they inspire; and Empathy, individualised attention to users' particular circumstances. The model's central claim is that quality is perceived relationally, and its operational consequence is that expectations and perceptions must be measured separately. A study by Juan and Nair (2022b) reveals that students' expectations are significantly higher than their perceptions of service quality of the university libraries in China.

In library settings, the model has proven durable, though not uncontested. Alam and Mezbah-ul-Islam (2020) demonstrated its adaptability to academic library assessment, and Butt et al. (2023) applied it successfully to web-based library services in Pakistan. Ghaedi et al. (2020) documented negative gaps across all five dimensions in Iranian medical university libraries, and Kekana and Kheswa (2020) reported a comparable pattern in South Africa. The consistency of these findings across very different systems is

striking and raises a question this study takes up: if negative gaps are near-universal, what distinguishes contexts is not their presence but their magnitude and their dimensional distribution.

Critics have argued that SERVQUAL's expectation component is conceptually unstable, in that users may report normative ideals rather than realistic anticipations, inflating gaps artificially. This critique has force, but it also points toward the analytic opportunity this study pursues. If expectation scores are shaped by what users take as their reference standard, then the identity of that standard becomes an empirical question worth asking directly.

Expectation–Confirmation Theory

Oliver's (1980) Expectation–Confirmation Theory supplies the mechanism SERVQUAL leaves implicit. Satisfaction, on this account, arises from the comparison of perceived performance against prior expectation. When performance meets or exceeds expectation, positive confirmation produces satisfaction; when it falls short, negative disconfirmation produces dissatisfaction. Crucially, the theory locates satisfaction in the relation between the two terms rather than in the absolute level of either. Churchill and Surprenant (1982) subsequently demonstrated that the relative contribution of expectation and performance varies by service type, a finding that matters for durable goods and infrastructure-heavy services in particular.

The theory has been under-exploited in library research, where it is more often cited than used. Yet it offers something the gap literature needs: an explanation of why gaps matter. A negative gap is not simply a measurement artefact; it is negative disconfirmation, and disconfirmation is the proximate cause of dissatisfaction. This study accordingly treats each dimensional gap as an instance of disconfirmation and asks the qualitative data to identify what generated the expectation that went unconfirmed.

Service Quality and Satisfaction in Academic Libraries

The empirical link between perceived library service quality and user satisfaction is well established. Choshaly and Mirabolghasemi (2019) confirmed it in Malaysia using structural equation modelling; Hasugian and Lubis (2021) in Indonesia; Filson and Barfi (2023) in Ghana. Within China, Chen et al. (2022) and Juan and Nair (2022a) found that weaknesses in emotional service and physical environment depressed satisfaction in non-state-owned university libraries, while Liu et al. (2024) established the centrality of digital usability and stable access to the contemporary user experience.

Which dimension matters most, however, is inconsistent across studies. Several report interpersonal dimensions as dominant: Davis et al. (2024) argue for empathy as the decisive factor in student engagement with library staff, and Zaineldeen et al. (2020) found personalised attention to be a leading predictor among Chinese students. Others locate primacy in the physical environment. Peng et al. (2022), comparing three universities in Wuhan, showed that spatial design, seating, and environmental comfort shaped satisfaction substantially; Kimaita et al. (2023) reached a similar conclusion in Kenyan higher education. This inconsistency is usually reported as a limitation of the literature. We read it differently: as evidence that dimensional primacy is a function of context, and specifically of whether basic infrastructural adequacy has been achieved. Where physical provision is sufficient, interpersonal quality differentiates; where it is strained, tangible conditions dominate. This proposition is testable, and the Chongqing setting offers a case in which infrastructure is under demonstrable demographic pressure.

Demographic Variation and the Present Study

Findings on demographic difference are similarly mixed. Kumar et al. (2022) reported gendered emphases in Indian universities, while Price and Fagan (2021) found students' perceptions of librarians to be more alike than different across gender and subject area. Alam et al. (2024) documented individual differences in Bangladeshi university libraries without finding them to be systematically demographic. The weight of recent evidence suggests that where service conditions are common to all users, evaluations converge, a proposition this study also examines.

The present study therefore extends the literature in a setting where the pressure on library infrastructure is acute, and it uses a mixed-methods design to move beyond the documentation of gaps toward their explanation.

METHOD

Design

An explanatory sequential mixed-methods design was employed. A quantitative strand established the magnitude and distribution of expectation–perception gaps and the predictive structure linking perceived quality to satisfaction; a subsequent qualitative strand explored how students accounted for those patterns. The design was chosen because the research questions are of two kinds, one asking whether and how much, the other asking why, and because the second could not be specified in advance of the first. Integration occurred at the interpretation stage through a joint display and the derivation of meta-inferences.

Setting and Participants

The study was conducted at five universities in Chongqing, China: Chongqing University, Southwest University, Chongqing Normal University, Chongqing Medical University, and Chongqing University of Posts and Telecommunications. These institutions were selected to span engineering, humanities, education, medicine, science, business, and technology, so that the sample would reflect varied disciplinary patterns of library use. Stratified random sampling was used, with strata defined by gender, faculty, and year of study. Following Krejcie and Morgan's (1970) table, a minimum of 384 respondents is required for a population exceeding 100,000 at a 95% confidence level; 400 students were recruited, approximately 80 from each university.

The sample comprised 220 male (55.0%) and 180 female (45.0%) students, distributed across the faculties of Arts (n = 94, 23.5%), Business (n = 104, 26.0%), Engineering (n = 97, 24.2%), and Science (n = 105, 26.2%), and across all four undergraduate years (Year 1, n = 108; Year 2, n = 88; Year 3, n = 100; Year 4, n = 104). Most respondents were regular library users: 29.2% attended daily and a further 35.8% three or four times weekly, with a mean of 10.1 hours per week spent in the library (SD = 3.89). This usage profile matters for interpretation, since it indicates that evaluations were grounded in sustained rather than occasional contact.

For the qualitative strand, semi-structured interviews were conducted with students drawn from the same five institutions, representing the Arts, Business, Science, Engineering, and Education faculties across undergraduate and postgraduate levels. Interviews lasted 15–20 minutes. Eight information-rich cases (P1–P8) were selected for detailed reporting; they were chosen because their accounts were analytically dense and collectively spanned the disciplinary and study-level range of the wider interview set.

Instruments

Three questionnaires and an interview protocol were used. Questionnaire A measured expectations of library service quality and Questionnaire B measured perceptions of actual service quality; both contained 30 items organised by SERVQUAL dimension: Tangibility (5 items), Reliability (7), Responsiveness (4), Assurance (6), and Empathy (8). Both were worded in parallel so that gap scores would be directly comparable. Questionnaire C measured satisfaction using 8 items addressing the usefulness of resources, the supportiveness of the learning environment, staff assistance, service effectiveness, and the contribution of library services to academic work. All items used a five-point Likert scale (1 = strongly disagree to 5 = strongly agree).

Items were adapted from validated SERVQUAL-derived instruments and reworded for the Chinese university library context. Sample items include "The library is well equipped with modern facilities" (Tangibility, perception) and "The library provides individual guidance, such as career planning and counselling" (Empathy, perception). The interview protocol comprised two core open questions, one on

perceived service quality and its most and least satisfying aspects, one on suggestions for improvement, each supported by prompts covering facilities, resources, staff support, and digital services.

Validity and Reliability

Three experts in education, information services, and research methodology reviewed the instruments for face validity, content validity, construct alignment, language clarity, and relevance. Mean expert ratings on a five-point scale were 4.67 for face validity, 5.00 for content validity, and 4.33 for construct validity. The interview protocol was rated on a ten-point scale, yielding means of 8.7 (face validity), 8.8 (content validity), and 8.7 (language clarity). Minor wording revisions followed.

A pilot study with 45 Chongqing university students, excluded from the main sample, established internal consistency. In the main study, Cronbach's alpha ranged from .808 to .861 across expectation subscales and from .812 to .857 across perception subscales, with the satisfaction scale reaching .889. Sampling adequacy was confirmed (KMO = .89) and Bartlett's test of sphericity was significant, $\chi^2(435) = 1250.34$, $p < .001$. Qualitative trustworthiness was addressed through the expert review of the protocol, verbatim transcription, iterative coding, and the retention of participant-attributed quotations to allow readers to assess the fit between evidence and interpretation.

Procedure and Analysis

Data were collected over approximately four weeks using self-administered questionnaires distributed in print and, where access conditions required, online. Interviews were conducted individually, face-to-face or online, and audio-recorded with consent.

Quantitative analysis used SPSS. Missing values did not exceed 2% on any variable and were addressed by mean substitution for continuous and mode substitution for categorical variables; no case was removed. All standardised residuals fell within ±3.29. Skewness and kurtosis for all composite variables fell within ±2, and graphical diagnostics (histogram, boxplot, Q–Q plot) supported approximate normality, justifying parametric procedures. Paired-samples t-tests examined expectation–perception differences, Pearson correlations examined bivariate associations, and standard multiple regression estimated the joint and unique contributions of the five perceived dimensions to satisfaction.

Qualitative data were analysed thematically following Braun and Clarke's (2006) six-phase framework. Transcripts were read repeatedly for familiarisation; initial codes were generated and collated into candidate themes; themes were reviewed for internal coherence and external distinctiveness, then defined and named. Coding for the improvement-focused question generated over 220 meaning units, reduced to 16 preliminary codes and clustered into four themes through constant comparison across faculties and study levels.

Ethical Considerations

Institutional ethical approval and permission from participating universities were obtained before data collection. Participation was voluntary and informed consent was secured. No personally identifying information was collected; demographic data were reported in aggregate. Participants were free to withdraw at any point without consequence. Data were stored in password-protected files accessible only to the researcher.

RESULTS

Descriptive Statistics

Table 1 reports composite means for the study variables. Expectation means clustered tightly near 4.00 with modest dispersion (SD ≈ 0.80), indicating that students held consistently high and broadly shared expectations. Perception means were lower (3.57–3.67) and more variable (SD ≈ 0.91–0.97), suggesting that

actual experience differed more between students than aspiration did. Satisfaction was moderate (M = 3.61, SD = 0.706). Figure 1 displays these means graphically, with error bars showing ±1 SD.

Table 1. Descriptive statistics for composite variables (N = 400)

Variable	M	SD	Skewness	Kurtosis
Expectation – Tangibility	3.98	0.807	0.036	−1.463
Expectation – Reliability	4.01	0.821	−0.014	−1.517
Expectation – Responsiveness	4.07	0.833	−0.127	−1.550
Expectation – Assurance	4.00	0.815	0.005	−1.494
Expectation – Empathy	3.99	0.813	0.018	−1.488
Perception – Tangibility	3.59	0.954	−0.034	−0.940
Perception – Reliability	3.66	0.913	−0.058	−0.850
Perception – Responsiveness	3.67	0.975	−0.120	−1.003
Perception – Assurance	3.63	0.952	−0.064	−0.943
Perception – Empathy	3.57	0.963	0.013	−0.974
Student satisfaction	3.61	0.706	0.016	−0.284

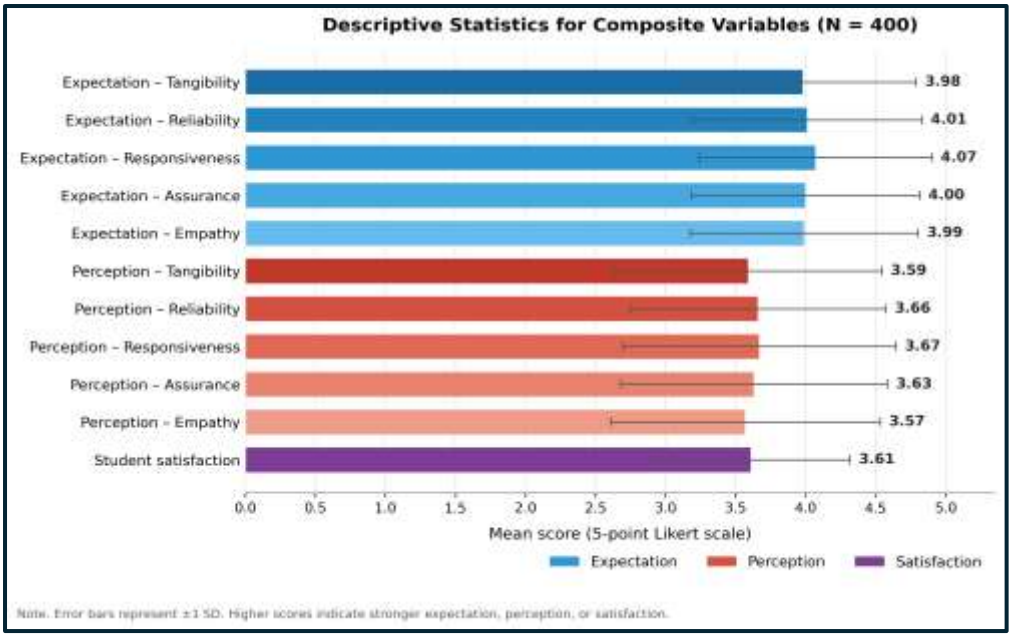

Figure 1. Mean expectation, perception, and satisfaction scores for the composite variables (N = 400). Error bars represent ±1 SD. All items were rated on a five-point Likert scale.

The asymmetry between the two blocks is itself informative. Homogeneous expectations paired with heterogeneous perceptions indicate that students arrive with a common standard and depart with divergent experiences, a pattern consistent with variable service delivery rather than variable aspiration.

Expectation–Perception Gaps

Overall expectation (M = 3.992, SD = 0.182) exceeded overall perception (M = 3.591, SD = 0.214). The paired-samples t-test confirmed a significant negative gap, mean difference = 0.401, 95% CI [0.391, 0.412], t(399) = 74.58, p < .001, Cohen's d = 3.73.

The magnitude of this effect requires comment. A d above 3 is far outside the range typically reported in service quality research and reflects the very small standard deviation of the difference score (SD = 0.108) relative to a substantively modest mean difference of 0.401 scale points. The finding should therefore be read as evidence of consistency rather than severity: nearly every student in the sample reported a shortfall, and reported one of similar size. The gap is a shared condition of the student population, not a concentrated complaint among a dissatisfied minority, which is precisely the pattern one would expect if it originates in institutional service conditions rather than in individual experience. Dimension-level results are reported in Table 2. Negative gaps were significant across all five dimensions, ranging from 0.394 (Reliability) to 0.406 (Assurance).

Table 2. Paired-samples t-tests for SERVQUAL dimensions (N = 400)

Dimension	Expectation M (SD)	Perception M (SD)	Gap	t(399)	p	d
Tangibility	3.979 (0.366)	3.575 (0.444)	0.404	34.05	< .001	1.70
Reliability	4.010 (0.306)	3.616 (0.368)	0.394	43.84	< .001	2.19
Responsiveness	3.993 (0.398)	3.596 (0.450)	0.397	33.95	< .001	1.70
Assurance	3.989 (0.316)	3.583 (0.358)	0.406	41.24	< .001	2.06
Empathy	3.989 (0.316)	3.583 (0.358)	0.406	41.24	< .001	2.06

What is notable is the flatness of the gap profile. The five dimensions differ by barely one hundredth of a scale point, which means that no single dimension accounts for the shortfall. Students did not report a library that excelled interpersonally while failing materially, or vice versa; they reported uniform underperformance relative to a uniform standard. This pattern constrains interpretation in a useful way, as the qualitative findings make clear.

Correlations Between Perceived Service Quality and Satisfaction

All five perceived dimensions correlated positively and significantly with satisfaction (Table 3), and overall perceived service quality correlated at r = .579, p < .001. Inter-dimensional correlations were very low (−.071 to .077).

Table 3. Correlations between perceived SERVQUAL dimensions and student satisfaction (N = 400)

Variable	r with satisfaction	p
Overall perceived service quality	.579	< .001
Tangibility	.309	< .001
Reliability	.290	< .001
Empathy	.267	< .001
Assurance	.241	< .001
Responsiveness	.235	< .001

The near-orthogonality of the five dimensions is unusual for SERVQUAL data, where substantial inter-correlation is the norm, and it warrants caution: it may reflect genuine dimensional independence in this setting, or it may indicate that respondents evaluated each block of items with limited reference to the others. Either way, it has an analytic consequence: the dimensions contribute largely non-overlapping variance to the regression model reported below, so their standardised coefficients can be read as relatively clean estimates of unique contribution.

Prediction of Student Satisfaction

Regression diagnostics supported the analysis: Durbin–Watson was 2.035, tolerance values ranged from .983 to .997, and VIF values from 1.003 to 1.018. The model was significant, $F(5, 394) = 41.51$, $p < .001$, explaining 34.5% of the variance in satisfaction ($R^2 = .345$, adjusted $R^2 = .337$). All five dimensions were significant positive predictors (Table 4).

Table 4. Multiple regression predicting student satisfaction (N = 400)

Predictor	B	SE	β	t	p
Constant	0.256	0.237	n/a	1.08	.281
Tangibility	0.194	0.026	.304	7.41	< .001
Reliability	0.205	0.031	.266	6.52	< .001
Responsiveness	0.137	0.026	.218	5.32	< .001
Assurance	0.167	0.033	.211	5.14	< .001
Empathy	0.233	0.035	.269	6.58	< .001

Tangibility carried the largest standardised weight (β = .304), followed by Empathy (β = .269), Reliability (β = .266), Responsiveness (β = .218), and Assurance (β = .211). Two features of this ordering deserve emphasis. First, the spread between the strongest and weakest predictor is modest, which argues against any strategy that concentrates resources on one dimension. Second, the two highest coefficients belong to dimensions that are conceptually opposed (the material and the interpersonal), suggesting that satisfaction is not produced along a single axis of service improvement. That 34.5% of variance is explained by library service quality alone is substantial given that student satisfaction is also shaped by teaching quality, workload, peer relationships, and institutional climate. It positions the library as a material contributor to the student experience rather than a peripheral support function.

Students' Perceptions of Service Quality

Thematic analysis of the interview data produced five themes (Responsiveness, Tangibility, Assurance, Empathy, and Reliability) corresponding to the SERVQUAL dimensions. The correspondence was not imposed analytically; the coding proceeded inductively and the dimensions emerged as the organising structure of students' own accounts.

Responsiveness was the most frequently raised theme. Students construed it relationally rather than temporally: what mattered was acknowledgement as much as speed. As P1 put it, being greeted and asked whether help was needed produced a feeling of being respected rather than overlooked. Digital responsiveness featured prominently, with P3 describing an online chat exchange answered within minutes and accompanied by explanatory links. Delays were tolerated where staff communicated; silence was not.

Tangibility was described as evidence of institutional priority. Students linked physical conditions to concentration and stamina, citing lighting and temperature that permitted long study sessions (P3), group discussion rooms that made the library a place for collaboration rather than only silent work (P2), while identifying specific deficits, most persistently the scarcity of charging outlets (P4) and slow computers (P1).

Assurance operated as a precondition for help-seeking. Confident, step-by-step guidance through database searches reduced anxiety about unfamiliar systems (P2, P3), whereas uncertainty on the part of newer staff produced confusion and, by implication, reluctance to ask again (P1). This was reported as particularly consequential for first-year students.

Empathy was characterised as the dimension that made the library socially accessible. Students described patient explanation as licensing further questions (P5) and the absence of judgment about basic queries as making the library feel safe (P1). Notably, empathetic interaction was described as compensating for infrastructural shortcoming, a claim the quantitative data neither confirm nor contradict, but which bears on how libraries under resource constraint might prioritise.

Reliability was noticed chiefly in its breach. A catalogue showing a book as available when it was not (P1) wasted time in a way students recalled vividly; conversely, stable database access was described as unremarkable but essential (P3). Communication during outages was identified as the factor that restored trust (P4). Reliability, in students' accounts, functioned as infrastructure in the strict sense: invisible when working, decisive when not.

Students' Improvement Priorities

Four themes emerged from the improvement-focused analysis. **Facilities** was the most frequently raised (approximately 62% of respondents), concerning overcrowding at peak periods, insufficient quiet areas, inadequate electrical outlets, uneven ventilation, and the absence of dedicated postgraduate research space. Proposed remedies included seat-reservation systems, ergonomic furniture, extended opening hours, and noise-controlled zones.

Resources followed (58%), centring on the currency of collections and the breadth of database subscriptions. Students in research-intensive disciplines were most emphatic: a postgraduate engineering student reported finding no journals from their field published in the previous two years, and a business student noted the absence of the data platforms their coursework presupposed. Several proposed formal mechanisms for student input into acquisition decisions.

Staff attitude and training was raised by 46% of respondents. Students distinguished courtesy, which they generally acknowledged, from competence in advanced database navigation, which they found uneven. One engineering student observed that students sometimes knew more than staff about research software. Suggestions focused on regular professional development and a feedback mechanism for service quality.

Digital support (39%) concerned system stability, interface design, and the availability of assistance outside staffed hours. Reported problems included overnight catalogue crashes, difficult navigation, and desynchronised overdue notices. Proposals included a 24-hour chatbot, interface redesign, and recorded tutorials for students unable to attend live workshops.

Integration

Table 5 presents the joint display linking quantitative findings, qualitative findings, and the meta-inferences drawn from their intersection.

Table 5. Joint display of quantitative and qualitative findings

Quantitative finding	Corresponding qualitative finding	Meta-inference

Significant negative gaps across all five dimensions, of near-identical magnitude	Concerns raised across facilities, resources, staff competence, responsiveness, and digital systems, without any dimension dominating	The shortfall is systemic rather than localised; it reflects a general standard applied across the service, not a specific failure
Tangibility strongest predictor (β = .304)	Physical conditions described as evidence of institutional priority; congestion and outlet scarcity most frequently cited	Under demographic pressure, tangible conditions carry symbolic as well as functional weight, amplifying their effect on satisfaction
Empathy second strongest (β = .269)	Empathetic interaction described as compensating for infrastructural shortcoming	Interpersonal quality functions as a partial substitute for material provision, offering a lower-cost improvement pathway
Reliability significant (β = .266); near-zero inter-dimensional correlation	Reliability noticed only when it fails; outage communication restores trust	Reliability operates as a threshold condition rather than a continuous contributor, consistent with its statistical independence from other dimensions
Model explains 34.5% of satisfaction variance	Students describe library conditions as shaping study efficiency, concentration, and sense of belonging	Library service quality is constitutive of the academic experience, not adjunct to it

The integration yields one finding that neither strand produces alone. The quantitative data establish that expectations were high, homogeneous, and unmet across every dimension; the qualitative data reveal the source of that homogeneity. Students consistently invoked external reference points commercial digital platforms whose interfaces set their standard for navigation, and better-resourced institutions whose database provision set their standard for collections. Expectations were not calibrated to their own library's trajectory. This explains the otherwise puzzling combination of documented institutional investment and persistent negative disconfirmation: improvement measured against an institution's own past is invisible to students measuring against an external present.

DISCUSSION

Gaps as Negative Disconfirmation

The finding of significant negative gaps across all five dimensions is consistent with SERVQUAL applications in academic libraries internationally (Ghaedi et al., 2020; Kekana & Kheswa, 2020; Mahmood et al., 2023). What the present study adds is an account of the uniformity of the gap profile, which the flat dimensional distribution in Table 2 makes conspicuous.

Read through Expectation–Confirmation Theory, this uniformity is diagnostic. Oliver (1980) holds that satisfaction follows from the comparison of performance against a prior standard; if the standard were generated by experience within a given library, one would expect dimension-specific variation reflecting that library's particular strengths and weaknesses. The observed flatness suggests a standard imported wholesale instead. The interview data support this reading directly: students judged navigation against commercial platforms and collections against better-resourced universities. Expectation, in this setting, is externally anchored.

The implication for theory is that expectation should be treated as a variable with identifiable sources rather than as a fixed baseline. This bears on the standing critique that SERVQUAL's expectation component measures ideals rather than realistic anticipations. Our findings suggest the distinction may be less consequential than assumed: if students genuinely use external services as their comparison standard, then what looks like an inflated ideal is in fact a realistic anticipation formed elsewhere. The practical

consequence is uncomfortable but important: libraries cannot close gaps by improving relative to their own history, because that is not the comparison being made.

The primacy of Tangibility

That Tangibility emerged as the strongest predictor of satisfaction diverges from studies identifying interpersonal dimensions as dominant (Davis et al., 2024; Zaineldeen et al., 2020) and converges with those emphasising the physical environment (Kimaita et al., 2023; Peng et al., 2022). We proposed in Section 2.3 that this inconsistency is not noise but a context effect, and the present findings are consistent with that proposition.

The mechanism appears to be one of threshold. Where physical provision is adequate, it recedes from evaluative attention and interpersonal quality differentiates between institutions. Where provision is strained, as it is in Chongqing (where student numbers have outpaced library capacity), tangible conditions become salient daily and dominate evaluation. The qualitative data show students encountering this strain concretely: time lost searching floors for a seat, sitting on the floor to reach a power outlet, discomfort during extended sessions in summer. These are not aesthetic complaints but impediments to the activity the library exists to support.

A second mechanism operates alongside the functional one. Students explicitly interpreted physical conditions as evidence of institutional priority, reading inadequate maintenance as indifference to student welfare. Tangibility thus carries symbolic weight, and its effect on satisfaction is amplified accordingly. This interpretation aligns with Peng et al. (2022) on spatial design as communicative, and it suggests that the frequent assumption that digital transformation diminishes the importance of physical facilities is mistaken. Digital study depends on physical infrastructure such as power, connectivity, workstations, and ergonomic seating, so investment in digital resources without corresponding attention to the spaces in which those resources are used is likely to underdeliver on satisfaction.

Empathy as Compensatory

Empathy's position as the second strongest predictor, and specifically the qualitative finding that empathetic interaction compensated for infrastructural shortcoming, has practical significance for resource-constrained institutions. Where capital investment in facilities operates on multi-year cycles, staff development operates on much shorter ones. Almeida and Tidal (2022) and Davis et al. (2024) have argued for empathy as a professional competence rather than a personal disposition; the present findings support that framing and give it an economic rationale.

The compensation claim should be stated carefully. Our data show that students described empathetic interaction as offsetting material limitation; they do not establish that satisfaction is statistically compensated in this way, which would require an interaction term the present design cannot support. This is a limitation and a clear direction for subsequent work: a moderation analysis testing whether Empathy's effect on satisfaction strengthens as perceived Tangibility declines would convert a reported perception into a testable mechanism.

Reliability as a Threshold Condition

Reliability behaved distinctively in both strands. Quantitatively, it was a significant predictor while remaining almost uncorrelated with the other dimensions. Qualitatively, students described noticing it only in failure. Together these suggest that Reliability operates as a threshold rather than a continuous contributor: below a certain level of dependability, satisfaction is impaired; above it, further reliability yields diminishing returns. This reading is consistent with Liu et al. (2024) on the centrality of stable access to digital library experience, and it has an operational implication. Reliability should be managed to a standard rather than optimised toward a maximum, with resources redirected to dimensions where returns remain linear once that standard is met. It also indicates the value of investment in outage communication, which students identified as the mechanism by which trust is restored after failure.

Theoretical Implications

The study makes three theoretical contributions. First, it specifies a context-dependency in SERVQUAL's dimensional structure: the relative weight of dimensions varies with infrastructural adequacy, such that tangible conditions dominate where capacity is strained and interpersonal conditions differentiate where it is not. This converts an inconsistency in the literature into a conditional proposition available for testing elsewhere.

Second, it advances the use of Expectation–Confirmation Theory in library research from citation to mechanism. By identifying the external anchoring of expectation, the study explains why negative disconfirmation persists through periods of documented improvement, a pattern that gap statistics alone render puzzling.

Third, it demonstrates the diagnostic surplus of gap-based over perception-only measurement. Perception means near 3.6 would, in a perception-only design, read as moderate and unremarkable. Placed against expectations near 4.0, the same values identify a systemic and universally experienced shortfall. The two designs support different conclusions from the same underlying performance, and library assessment practice should reflect this.

Practical Implications

Four recommendations follow, ordered by the evidence rather than by ease of implementation.

Tangibility warrants first priority, given both its predictive weight and the specificity of student accounts. The immediate targets are capacity management during peak periods (seat reservation systems, extended hours, overflow spaces), power provision, ventilation, and the replacement of ageing workstations. These are unglamorous interventions with a strong evidentiary warrant.

Empathy warrants near-term investment because it is achievable within budget cycles that facilities improvements are not. Training should treat empathetic practice as a competence (how to receive a basic question without conveying impatience, how to sustain help-seeking among first-year students) rather than as an attribute of individual temperament.

Reliability should be managed to a defined standard, with catalogue accuracy audits, monitored uptime for digital platforms, and, critically, a protocol for communicating outages. Students' accounts indicate that communication during failure substantially mitigates its effect on trust.

Resource currency requires structural rather than incremental attention. The complaint that core disciplinary material is a decade old is not resolved by marginal acquisition budgets; it points toward consortium purchasing across Chongqing institutions and transparent mechanisms for student input into acquisition. Given the regional concentration of universities, shared subscription arrangements are a plausible response to a problem individual libraries cannot solve alone.

LIMITATIONS AND FUTURE RESEARCH

Four limitations qualify these findings. First, the study is confined to five universities in a single municipality; while this focus is deliberate and addresses a documented geographic gap, it constrains generalisation, and the finding that Tangibility dominates should be understood as conditional on the infrastructural pressure characterising this setting rather than as a general ranking.

Second, the cross-sectional design captures expectations and perceptions at a single point. Expectations plausibly shift across the academic year, rising before examination periods when demand for space peaks, and a longitudinal design would establish whether gaps are stable or cyclical.

Third, all measures are self-reported and collected concurrently, so common method variance cannot be

excluded. Subsequent work might pair perceptual measures with behavioural data such as gate counts, loan records, or database session logs.

Fourth, the very large effect size for the overall gap reflects an unusually small standard deviation of difference scores. We have interpreted this as evidence of consistency across respondents, but replication in other settings is needed before the pattern can be treated as characteristic rather than sample-specific.

Beyond addressing these limitations, three directions appear productive: testing the proposed moderation of Empathy's effect by perceived Tangibility; comparative work across institutions differing systematically in infrastructural adequacy, which would test the threshold proposition directly; and research on expectation formation itself, tracing how students acquire the external standards against which they judge library services.

CONCLUSION

Students at five Chongqing universities expect more of their libraries than they perceive themselves to receive, and they do so consistently, across every dimension of service quality, and to a similar degree. Perceived quality nevertheless matters: the five SERVQUAL dimensions jointly explain a third of the variance in satisfaction, with physical conditions carrying the greatest weight and interpersonal support close behind.

The study's central argument is that these gaps are not adequately understood as internal performance deficits. Students judge their libraries against standards formed elsewhere, namely commercial digital platforms and better-resourced institutions, and improvement measured against an institution's own past therefore remains invisible to them. This has an uncomfortable implication for library management: closing gaps requires meeting an external standard, not merely exceeding a previous one. It also has a constructive one. Because expectations are anchored in specific, identifiable comparisons, they can be investigated, and libraries that understand what their students are comparing them to are better positioned to decide where investment will register.

REFERENCES

1. Alam, M. J., & Mezbah-ul-Islam, M. (2020). Service quality assessment model for academic libraries. Global Knowledge, Memory and Communication, 69(3), 181–198. https://doi.org/10.1108/GKMC-03-2020-0027
2. Alam, M. J., Matin, W.-U., & Begum, B. (2024). Individual differences in service quality of university libraries in Bangladesh. IFLA Journal, 50(1), 34–48.
3. Almeida, N., & Tidal, J. (2022). Library wayfinding and ESOL students: Communication challenges and empathy-based intervention. portal: Libraries and the Academy, 22(3), 453–474.
4. Appleton, L. (2020). Academic libraries and student engagement: A literature review. New Review of Academic Librarianship, 26(2), 189–213.
5. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.
6. Butt, K., Arif, M., & Rafiq, M. (2023). Assessing university students' satisfaction with web-based library services based on SERVQUAL model: A case of a Pakistani university. Libri, 73(2), 153–165.
7. Chen, L., Wang, J., & Xu, H. (2022). Service quality improvement in non-state-owned academic libraries in China: An empirical study. Library Management, 43(1/2), 32–47. https://doi.org/10.1108/LM-05-2021-0058
8. Choshaly, S. H., & Mirabolghasemi, M. (2019). Using SEM-PLS to assess users satisfaction of library service quality: Evidence from Malaysia. Library Management, 40(8/9), 556–570.
9. Churchill, G. A., & Surprenant, C. (1982). An investigation into the determinants of customer satisfaction. Journal of Marketing Research, 19(4), 491–504.

10. Davis, L., Albro, M., Mack, T., & Royse, M. (2024). Authentic connection: Engaging with students through empathy. portal: Libraries and the Academy, 24(1), 45–62.
11. De Groote, S. D., & Scoulas, J. (2022). The impact of the academic library on students' success, in their own words. portal: Libraries and the Academy, 22(2), 355–374.
12. Du, X., & Zhu, Y. (2021). Regional disparities in digital resource allocation among Chinese academic libraries. The Electronic Library, 39(6), 1248–1265. https://doi.org/10.1108/EL-02-2021-0048
13. Eshbach, B. E. (2020). Supporting and engaging students through academic library programming. The Journal of Academic Librarianship, 46(3), 102129.
14. Filson, C., & Barfi, K. A. (2023). Assessing the quality of services at an academic library. Heliyon, 9(5), e22449.
15. Ghaedi, R., Valizadeh-Haghi, S., Ahmadi, E., Zeraatkar, Z., & Baghestani, A. (2020). Gaps between users' expectations and their perceptions on service quality of college libraries. DESIDOC Journal of Library & Information Technology, 40(2), 479–484.
16. Gyau, E. B., Liu, J., & Kwakye, O. (2021). Evaluation of user satisfaction with academic libraries services based on students' perspectives. OALib Journal, 8, 5–17.
17. Hasugian, J., & Lubis, D. (2021). Library service quality and student trust: A case study of the University of Sumatera Utara Library, Indonesia. DESIDOC Journal of Library & Information Technology, 41(5), 420–429.
18. Juan, L., & Nair, S. M. (2022a). Service quality and students' satisfaction in higher vocational college in China. International Journal of Arts Humanities and Social Sciences Studies, 7(5), 36-46.
19. Juan, L. & Nair, S .M. (2022b). Students' expectation and perception on service quality in a public Higher Vocational College in China, International Academic Journal of Education & Literature, 3(3), 1-3.
20. Kekana, M. D., & Kheswa, S. (2020). The gap between user perceptions and expectations of students at the main library of the University of KwaZulu-Natal. SA Journal of Information Management, 22(1), a1131.
21. Kimaita, J. N., Kabue, H., Ochola, P., & Kiriinya, S. N. (2023). Influence of service tangibility on student satisfaction in schools of business of chartered universities in Kenya. International Journal of Research and Innovation in Social Science, 7(7), 53–59.
22. Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30(3), 607–610.
23. Kumar, V., Singh, P., & Ahmad, M. (2022). Gender differences in perception of academic library services: Evidence from Indian universities. Library Philosophy and Practice, Article 6894.
24. Liu, X., Zhang, Y., & Huang, W. (2024). Enhancing digital library usability: An empirical study of Chinese academic libraries. The Journal of Academic Librarianship, 50(1), 102700. https://doi.org/10.1016/j.acalib.2023.102700
25. Mahmood, K., Rehman, S., & Ashiq, M. (2023). The users' expectations of library service quality in the college libraries of Pakistan: A LibQUAL study. Performance Measurement and Metrics, 24(1), 14–30.
26. Malik, S. A., Fatima, T., Jia, Y., & Pannu, H. (2023). The influence of library service quality, library image, and trust on loyalty: The mediating role of perceived service value and satisfaction. International Journal of Quality & Reliability Management, 40(3), 234–256.
27. Mendoza, R. (2024). Digital readiness and student satisfaction with academic library services during and after COVID-19. Library & Information Science Research, 46(1), 101223. https://doi.org/10.1016/j.lisr.2023.101223
28. Murphy, J. E., Premji, Z., & Jones, R. (2021). A scoping review of personal librarian programs in academic libraries. College & Undergraduate Libraries, 29(1), 23–53.
29. Oh, D. (2022). How to measure service quality, customer satisfaction and loyalty of public library users. Journal of Librarianship and Information Science, 54(1), 1–18.
30. Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. Journal of Marketing Research, 17(4), 460–469.

31. 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.
32. Peng, L., Wei, W., Fan, W., Jin, S., & Liu, Y. (2022). Student experience and satisfaction in academic libraries: A comparative study among three universities in Wuhan. Buildings, 12(5), 1–19.
33. Price, E., & Fagan, J. (2021). More alike than different: Student perceptions of academic librarians by genders and subject areas. The Reference Librarian, 62(3–4), 221–263.
34. Qu, M. (2023). Exploring patron behavior in an academic library: A Wi-Fi-connection data analysis. Education and Information Technologies, 29(4), 11235–11256.
35. Zaineldeen, S., Hongbo, L., & Koffi, A. L. (2020). Service quality dimensions and student satisfaction in Chinese universities. International Journal of Higher Education, 9(5), 1–12.