

AI Customer Service, Perceived Diagnosticity, and Brand Loyalty in Cosmetics Live-Streaming Commerce

Weixiang Gan¹, Mengfei Xiao¹, Shashi Kant Gupta^{1,2*}

¹ Lincoln University College, Petaling Jaya, Selangor Darul Ehsan-47301, Malaysia

^{2*} Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India.

pdf.weixiang@lincoln.edu.my, pdf.Mengfei@lincoln.edu.my, shashigupta@lincoln.edu.my
(Corresponding author)

Abstract. In China's cosmetics live-streaming commerce, AI customer service is evolving from a tool for automated replies into a digital service touchpoint that shapes product assessments, service evaluations, and brand relationships. Because cosmetics are experiential products that carry pronounced product-fit uncertainty, consumers attend not only to response speed but also to whether AI customer service helps them assess a product's fit for their skin type, functional needs, shade preference, and usage context. Drawing on the stimulus-organism-response (S-O-R) framework and uncertainty-reduction theory, this study develops a serial-mediation model linking AI customer service experience, perceived diagnosticity, trust, satisfaction, and brand loyalty, and tests it with partial least squares structural equation modelling (PLS-SEM) using survey data from 450 Chinese consumers of cosmetics live-streaming commerce. The results indicate that AI customer service experience is positively associated with perceived diagnosticity, which is in turn associated with greater trust in the brand and its service information. Trust is associated with brand loyalty both directly and indirectly through satisfaction, and the bias-corrected bootstrap estimates are consistent with sequential mediation through perceived diagnosticity, trust, and satisfaction. These findings suggest that, by supplying tailored decision support that reduces product-fit uncertainty, AI customer service can help convert brief service interactions into stronger brand preference. The study positions AI customer service as an integral element of the digital service systems that underpin live-commerce platforms and offers operational guidance for deploying it as a decision-support mechanism rather than a mere cost-saving channel.

Keywords: AI customer service; live-streaming commerce; cosmetics; brand loyalty; perceived diagnosticity; digital service systems

1. Introduction

Live-streaming commerce has become an important setting in which consumers obtain product information, conduct interactive consultation, and make immediate purchases. This is especially true in experiential product categories such as apparel and cosmetics, where consumers often find it challenging to assess whether a product really fits them via static product pages alone. Live-streaming rooms can therefore support purchase decisions by reducing uncertainty through demonstration, explanation and interaction. Previous research has shown that, in apparel and cosmetics live streaming, the vicarious product trials and real-time interaction of streamers help consumers develop trust and reduce product uncertainty, thus increasing purchase intention (Lu & Chen, 2021). Research on live streaming shows that interactivity, social presence, instant feedback and dialogic quality jointly influence trust, engagement and purchase responses (Jiang et al., 2022; Konya-Baumbach et al., 2023; Cui et al., 2024). Compared with conventional image-and-text e-commerce, the advantage of live streaming is not merely that it is more engaging, but that it integrates product information, social interaction and decision support into a single transactional scene.

In this study, AI customer service denotes the brand- or store-deployed automated conversational agents that consumers engage with for pre-sale and in-sale consultation during live-streaming shopping. These agents include rule-based and generative chatbots, platform and brand-store service bots, and human–AI hybrid arrangements in which routine queries are answered automatically while complex cases are escalated to human agents. This conceptualisation deliberately centres on the service function performed rather than on any single underlying technology, so that the findings speak to AI customer service as a service mechanism instead of to a specific chatbot implementation.

However, existing live-streaming commerce research has mainly focused on streamer characteristics, interactive atmosphere, social presence, and impulse buying, while paying insufficient attention to AI customer service inside live-streaming rooms. Recent studies have begun discussing virtual streamers, AI streamers and AI live-streaming agents, but the dominant questions still tend to centre on identity, anthropomorphic cues, language naturalness and purchase intention, instead of treating AI customer service as an independent service touchpoint that can explain long-term relational outcomes (Gao et al., 2023; Zhang et al., 2026). Meanwhile, AI customer service research has demonstrated that service quality can affect customer loyalty through perceived value, cognitive trust, affective trust, and satisfaction, but its main empirical settings are still general e-commerce or digital service environments (Chen et al., 2023; Hsu & Lin, 2023; Huang et al., 2024; Rita et al., 2026). In other words, although both streams of research have advanced rapidly, they have not adequately addressed a question closer to commercial reality: when consumers interact with AI customer service during cosmetics live-streaming shopping, can such interaction be translated into long-term loyalty toward the brand?

This question is important because cosmetics consumption is not a one-time price comparison. Shade, skin type, ingredients, efficacy and usage method are important individual differences. What consumers need is not just “someone responds”, but whether those responses help them judge more accurately the suitability of a product for themselves. In the social-commerce and live-streaming literature, information diagnosticity is regarded as an important cognitive mechanism that turns visibility, interactivity and product cues into purchase judgment, particularly for products with high uncertainty (Shin et al., 2013; Lu & Chen, 2021). Therefore, this study contends that the value of AI customer service in cosmetics live streaming should not be interpreted simply in terms of response speed or anthropomorphic quality. Whether it increases consumers’ perceived diagnosticity, that is, their capacity to judge more clearly how well a product fits their own needs, should provide a further explanation.

Based on this, the present paper makes three contributions. First, it extends AI customer service research from general digital service settings to the more contextually demanding setting of cosmetics

live-streaming commerce, which is characterised by high experientiality and high uncertainty. Second, this study introduces perceived diagnosticity into the relationship between AI customer service experience and brand loyalty to explain how service experience is accumulated into a brand relationship through a decision-support mechanism. Third, it explicitly differentiates “continued use of AI customer service itself” from “loyalty toward the brand”, thus avoiding the confounding of technology-adoption outcomes with relationship-marketing outcomes.

Framing the study within service science, we treat AI customer service as a component of the digital service systems that operate live-commerce platforms. Beyond answering isolated questions, brand- and platform-deployed service agents participate in pre-sale consultation, real-time processing of customer information, order- and after-sales coordination, and the routing of complex or sensitive cases to human agents. Locating AI customer service within these service workflows connects the present analysis to the operational concerns of digital service management and clarifies how a frontline service technology, rather than merely reducing response costs, can create relational value within the broader platform service ecosystem.

2. Theoretical Foundations and Research Hypotheses

2.1. AI Customer Service Experience and Brand Loyalty

The service-quality–loyalty chain proposes that customers’ positive evaluation of a service process is gradually transformed into trust, satisfaction, and repeat choice (Parasuraman et al., 1988; Morgan & Hunt, 1994; Oliver, 1999). AI customer service is not merely an information-retrieval tool, but a frontline touchpoint at which the brand presents itself to consumers in digital environments. Previous studies have revealed the role of AI customer service quality in increasing customer loyalty via perceived value, cognitive trust, affective trust, and satisfaction. Other studies have demonstrated that the quality of service recovery and the conversational quality of chatbots significantly affect satisfaction (Chen et al., 2023; Hsu & Lin, 2023). Consistent with this, recent service-science evidence indicates that chatbot responsiveness, ease of use, and personalization jointly enhance the customer experience in e-commerce (Megdadi et al., 2025; Hoang & Tran, 2026). In more specific service contexts, learning-based trust, emotional expression, and light human–AI collaboration can also shape the effects of AI customer service on satisfaction and repatronage outcomes (Huang et al., 2024; Zhang et al., 2024). This implies that when AI customer service is clear, timely and need-relevant, consumers are more likely to attribute this high-quality contact to the overall capability of the brand, thus forming more positive relational judgements.

Accordingly, H1 is proposed: AI customer service experience positively affects brand loyalty.

2.2. AI Customer Service Experience and Perceived Diagnosticity

Perceived diagnosticity emphasises whether information helps consumers make more accurate and more confident judgements. In cosmetics live streaming, consumers’ most important question is not whether information exists, but whether the information can help them judge whether a product is suitable for them. Prior research has identified the effects of product-related risk, information quality and interaction cues on consumers’ evaluation of information diagnosticity (Shin et al., 2013; Wang & Chang, 2013; Lu & Chen, 2021). Compared with standardised live explanations, AI customer service can provide more targeted responses based on a consumer’s skin type, shade preference, budget and specific questions. It may thus turn general information in a live-streaming room into diagnostic information that is more useful to the individual consumer.

Theoretically, perceived diagnosticity is positioned here as an organismic (O) state rather than as a property of the stimulus. Diagnosticity is not an objective attribute of the information delivered by AI customer service; it is a consumer’s subjective appraisal of the extent to which that information enables an accurate and confident judgement of product fit. Two consumers who receive the same response may

form different diagnosticity perceptions depending on their skin concerns, prior knowledge, and decision goals. Because it is the outcome of cognitive processing of the service stimulus, perceived diagnosticity belongs to the organism component of the S–O–R sequence, mediating between the AI customer service stimulus and consumers’ relational responses, rather than constituting an information-quality feature of the stimulus itself.

Accordingly, H2 is proposed: AI customer service experience positively affects perceived diagnosticity.

2.3. Perceived Diagnosticity, Trust, and Satisfaction

Research on uncertainty reduction in live-streaming commerce finds that trust is an important psychological bridge in the purchase of experiential products (Lu & Chen, 2021). Broader social-commerce studies have consistently proposed that interaction, social presence, information quality and responsiveness increase trust and then influence purchase, continued purchase and word-of-mouth behaviours (Hajli et al., 2017; Liang et al., 2019; Meilatinova, 2021). Consumers tend to rate a source as reliable and trustworthy when they perceive that the source helps them understand a product more accurately. In the field of cosmetics, if AI customer service can effectively respond to the question of “whether it suits me”, consumers are more likely to trust the specific advice in the current interaction and to transfer the perceived capability to their overall trust in the brand, consistent with a trust-transfer mechanism (Zhao et al., 2019). Satisfaction is also shaped by trust (Eni et al., 2024). When a brand is perceived to provide reliable information and useful advice in a high-uncertainty environment, consumers are more likely to believe that the interaction meets their expectations and to develop a positive evaluation of the overall service experience.

The path from AI customer service experience to satisfaction (H4) can be grounded more explicitly in service-quality and expectation-confirmation reasoning. Service-quality theory holds that favourable evaluations of a service encounter, particularly its responsiveness, reliability, and assurance, feed directly into overall satisfaction (Parasuraman et al., 1988; Hsu & Lin, 2023). Expectation-confirmation theory adds that satisfaction arises when perceived performance meets or exceeds prior expectations (Oliver, 1999). When AI customer service supplies clear, timely, and need-relevant assistance, it is likely to confirm or surpass consumers’ expectations of the interaction, so that a positive service experience is translated into satisfaction in addition to, and partly independently of, its effect through trust.

Accordingly, the following hypotheses are proposed:

H3: Perceived diagnosticity positively affects trust.

H4: AI customer service experience positively affects satisfaction.

H5: Trust positively affects satisfaction.

2.4. Trust, Satisfaction, and Brand Loyalty

Long-term relationships are built on trust and satisfaction, as relationship-marketing theory has long highlighted (Morgan & Hunt, 1994; Oliver, 1999). Both constructs have consistently been found to drive continued purchasing, repatronage and word-of-mouth communication in social-commerce and AI-service research (Chen et al., 2023; Huang et al., 2024; Meilatinova, 2021). When consumers feel that a brand is both able and willing to offer reliable help, they are more likely to buy again, give the brand priority and recommend it proactively. Satisfaction, as consumers’ overall judgment of prior experience, also directly influences repurchase and word-of-mouth.

Accordingly, the following hypotheses are proposed:

H6: Trust positively affects brand loyalty.

H7: Satisfaction positively affects brand loyalty.

2.5. Chain Mediation

Following the logic above, AI customer service experience may not only directly influence loyalty but also sequentially influence loyalty through perceived diagnosticity, trust and satisfaction. In other words, AI customer service is the first means by which customers understand the fit between a product and their own needs. This support for judgement then reinforces trust in the brand, generates greater satisfaction, and is finally converted into stronger brand loyalty.

Thus, the following hypothesis is proposed: H8. Perceived diagnosticity, trust, and satisfaction play a serial (chain) mediating role in the relationship between AI customer service experience and brand loyalty.

Integrating the above hypotheses within the stimulus–organism–response (S–O–R) framework yields the conceptual model shown in Figure 1. AI customer service experience serves as the external stimulus (S); perceived diagnosticity, trust, and satisfaction constitute the internal organismic states (O) that process and accumulate this stimulus; and brand loyalty represents the behavioural-relational response (R). The model specifies one direct path (H1) and a set of mediating paths (H2–H7) that together form the serial mediation expressed in H8.

We deliberately model the influence of perceived diagnosticity on satisfaction as operating through trust rather than as an additional direct path. Conceptually, diagnosticity is a cognitive appraisal of decision support, whereas satisfaction is an affective-evaluative response to the encounter; the S–O–R and uncertainty-reduction logic implies that the cognitive benefit of diagnosticity is converted into affect once it has consolidated into trust in the brand and its information. A direct diagnosticity–satisfaction path was therefore not hypothesised. We none the less acknowledge it as a plausible alternative specification and return to it when discussing the boundaries of the model.

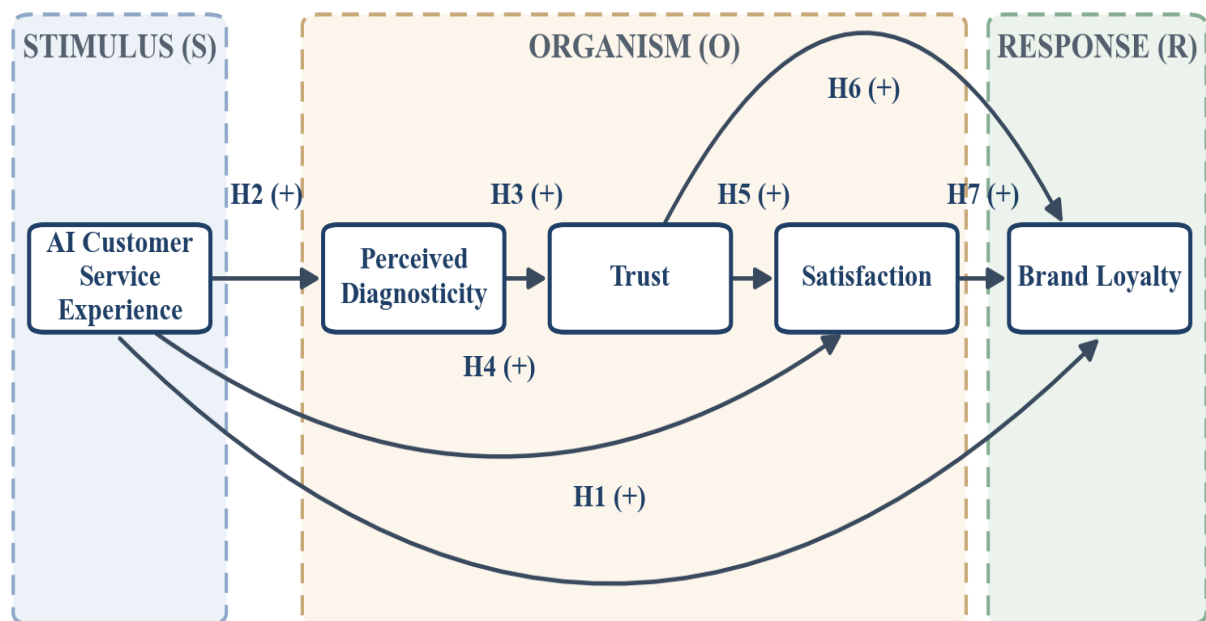


Fig.1: Conceptual framework based on the stimulus–organism–response (S–O–R) perspective.

Note. H1–H7 denote the hypothesised direct relationships (all expected to be positive); H8 denotes the hypothesised serial mediation path: AI customer service experience → perceived diagnosticity → trust → satisfaction → brand loyalty.

3. Research Design

3.1. Sample and Procedure

The study surveyed adult consumers of Chinese domestic cosmetics in the live-streaming commerce

context. Respondents were screened by three criteria: they had watched domestic cosmetics live streams in the previous three months; they had used a brand or store AI customer service during live-streaming shopping; and they could recall a specific AI customer service interaction. Respondents were recruited through a professional Chinese online survey platform (Credamo) between March and May 2025, combining platform-panel sampling with limited snowball referral across the major live-streaming platforms. Participation was voluntary and anonymous; respondents gave informed consent and received a small monetary incentive upon valid completion. To protect data quality, the survey accepted only one submission per device and IP address and embedded screening, attention-check, and logic-consistency items. After the screening described in Section 4.1, 450 valid responses were retained. The sample size meets the basic requirements of structural equation modelling. The mean age of respondents was 27.53 years ($SD = 5.73$). The sample comprised 337 female respondents (74.9%), 106 male respondents (23.6%) and 7 respondents who identified as other or preferred not to disclose their gender (1.6%). As for education, respondents with a bachelor's degree accounted for 60.4%, those with a junior-college diploma for 21.1%, and those with a master's degree or above for 11.8%. In terms of monthly disposable income, the RMB 3,000–5,999, RMB 6,000–9,999 and RMB 10,000–14,999 groups accounted for 26.7%, 30.2% and 24.0%, respectively. Regarding consumption experience, 303 respondents (67.3%) had previous experience with the focal brand. The most recent purchases were mainly in the skincare and colour-cosmetics categories, accounting for 47.3% and 32.4%, respectively. For distribution by platform, Douyin, Taobao Live, Kuaishou and Xiaohongshu accounted for 41.8%, 28.0%, 15.3% and 10.4%, respectively. Overall, the sample covers the major domestic live-streaming platforms, typical cosmetics categories, and consumer groups with differing characteristics. Although the sampling was non-probabilistic and the sample should not be read as statistically representative, it captures the principal consumer segments and provides a sound basis for examining consumers' self-reported perceptions, based on a recalled interaction, and their brand-relationship responses after engaging with AI customer service in cosmetics live-streaming shopping. The detailed sample profile is presented in Table 1.

Table 1. Profile of the Sample (N = 450)

Indicator	Category	Frequency	Percentage
Sample size	Valid responses	450	100.0%
Age	Mean = 27.53; SD = 5.73	n/a	n/a
Gender	Female	337	74.9%
	Male	106	23.6%
	Other / prefer not to disclose	7	1.6%
Education	Bachelor's degree	272	60.4%
	Junior-college diploma	95	21.1%
	Master's degree or above	53	11.8%
	Other	30	6.7%
Monthly disposable income	RMB 3,000–5,999	120	26.7%
	RMB 6,000–9,999	136	30.2%
	RMB 10,000–14,999	108	24.0%

Indicator	Category	Frequency	Percentage
	Other income ranges	86	19.1%
Prior brand-use experience	Yes	303	67.3%
	No / unclear	147	32.7%
Most recent purchase category	Skincare	213	47.3%
	Colour cosmetics	146	32.4%
	Other categories	91	20.2%
Live-streaming platform	Douyin	188	41.8%
	Taobao Live	126	28.0%
	Kuaishou	69	15.3%
	Xiaohongshu	47	10.4%
	Other platforms	20	4.4%

3.2. Construct Measurement

The questionnaire used a seven-point Likert scale. AI customer service experience was assessed mainly through five aspects: information clarity, response timeliness, problem-solving capability, need fit, and overall service quality. Perceived diagnosticity reflected the extent to which consumers perceived that, after interacting with AI customer service, they could judge product fit, reduce uncertainty, compare product differences, and make purchase decisions with greater confidence. Trust comprised consumers' perceptions of the reliability and authenticity of the information provided by the brand and its AI customer service, their general trust, and their willingness to rely on such information. Satisfaction consisted of overall satisfaction, expectation confirmation, interaction pleasantness, and retrospective evaluation. Brand loyalty captured consumers' future preference for the brand, repurchase intention, willingness to recommend, and continued preference.

All items were adapted from mature constructs drawn from prior studies of AI customer service, live-streaming commerce, perceived diagnosticity, trust, satisfaction, and brand loyalty. The wording was adapted to the Chinese domestic cosmetics live-streaming commerce context to ensure that the measurement accurately reflected consumers' self-reported perceptions and behavioural tendencies, recalled from a specific interaction with a brand or store AI customer service during live-streaming shopping. The construct sources are summarised in Table 2, and the full item wording is reported in Appendix Table A1.

Each construct was modelled reflectively. For AI customer service experience, the five facets (information clarity, response timeliness, problem-solving capability, need fit, and overall service quality) are treated as interchangeable manifestations of a single underlying evaluation of the AI service encounter rather than as distinct formative components: they are expected to covary, to share a common cause, and to be mutually substitutable without changing the meaning of the construct. Brand loyalty is likewise conceptualised as a unidimensional, primarily attitudinal loyalty intention (a general disposition toward continued preference, repurchase, and recommendation) rather than as a formative or higher-order construct, and its four indicators are treated as reflective manifestations of that disposition. This reflective specification is consistent with the high internal-consistency reliability and convergent validity reported in Section 4.3 and with the absence of indicator-level collinearity (all outer VIF < 2.1).

Table 2. Constructs, Representative Sources, and Application Context

Construct	Representative basis	Application context
AI customer service experience	Chen et al. (2023) Hsu & Lin (2023)	Integrates information quality, responsiveness, problem solving, and need fit
Perceived diagnosticity	Shin et al. (2013) Lu & Chen (2021)	Adapted to “whether it helps consumers judge cosmetics fit”
Trust	Morgan & Hunt (1994) Huang et al. (2024)	Reliability judgements about the brand and its AI service information
Satisfaction	Oliver (1999) Hsu & Lin (2023)	Overall evaluation after the current interaction
Brand loyalty	Oliver (1999) Chen et al. (2023)	Priority choice, repurchase, and recommendation

3.3. Analytical Strategy

The hypotheses were tested with partial least squares structural equation modelling (PLS-SEM) using SmartPLS 4. PLS-SEM was selected because the model is prediction-oriented, involves a complex serial-mediation structure, and does not require the multivariate-normality assumption of covariance-based SEM (Hair et al., 2019; Hair et al., 2022). Following established guidelines, a two-stage procedure was adopted. First, the measurement model was evaluated through indicator reliability and outer loadings, internal-consistency reliability (Cronbach’s α , the reliability coefficient ρ_A , and composite reliability), convergent validity (average variance extracted, AVE), and discriminant validity. Discriminant validity was assessed using both the Fornell–Larcker criterion and the heterotrait–monotrait ratio of correlations (HTMT), and the absence of indicator multicollinearity was confirmed through the variance inflation factor (VIF) (Henseler, Ringle, & Sarstedt, 2015; Dijkstra & Henseler, 2015). Second, after the measurement model met the recommended thresholds, the structural model was estimated by examining inner-model collinearity, the significance of the path coefficients, the coefficient of determination (R^2), the effect size (f^2), and the out-of-sample predictive relevance (Q^2). The significance of the direct and indirect (mediating) effects was assessed using bias-corrected bootstrapping with 5,000 subsamples (Preacher & Hayes, 2008; Zhao et al., 2010). Because the data were collected through a single-source self-report questionnaire, both procedural and statistical remedies were used to minimise common method bias, including anonymous responses, item-order randomisation, semantic differentiation among measures, Harman’s single-factor test, and a full-collinearity assessment (Podsakoff et al., 2003; MacKenzie & Podsakoff, 2012; Kock, 2015).

4. Results

4.1. Data Screening and Descriptive Statistics

Before model estimation, the raw data were screened to ensure data quality. A total of 540 questionnaires were initially collected. Cases that failed the eligibility (screening) questions, failed the embedded attention-check (instructed-response) items, exhibited straight-lining or insufficient response variance, were completed implausibly quickly (speeders), or were identified as multivariate outliers (Mahalanobis distance, $p < 0.001$) were removed sequentially, as summarised in Table 3. This procedure yielded 450 valid responses, corresponding to a retention rate of 83.3%. Because the questionnaire used a forced-response design, no missing values were present in the retained data set. Figure 2 visualises this stepwise screening cascade.

Two screening rules warrant clarification. The attention check was an instructed-response item embedded among the substantive measures (for example, “For quality-control purposes, please select ‘disagree’ for this item”); responses that did not give the instructed answer were treated as inattentive. Straight-lining was operationalised as identical answers across all items of at least one multi-item block (zero within-block variance). The 90-second speeding threshold reflects the length of the instrument: with 21 substantive items in addition to screening and demographic questions, a completion time below 90 seconds corresponds to fewer than roughly two seconds per item, which is implausibly fast for attentive reading and response. All thresholds were specified before data analysis.

Table 3. Summary of the Data-Screening Procedure

Step	Screening criterion	Cases removed	Remaining
0	Questionnaires initially collected	n/a	540
1	Failed eligibility / screening questions	47	493
2	Failed attention-check (instructed-response) items	18	475
3	Straight-lining / insufficient response variance	9	466
4	Speeders (completion time < 90 s)	10	456
5	Multivariate outliers (Mahalanobis D^2 , $p < 0.001$)	6	450
Final valid sample		n/a	450

Note. Retention rate = $450 / 540 = 83.3\%$. The questionnaire used a forced-response design; therefore no missing data were present after screening.

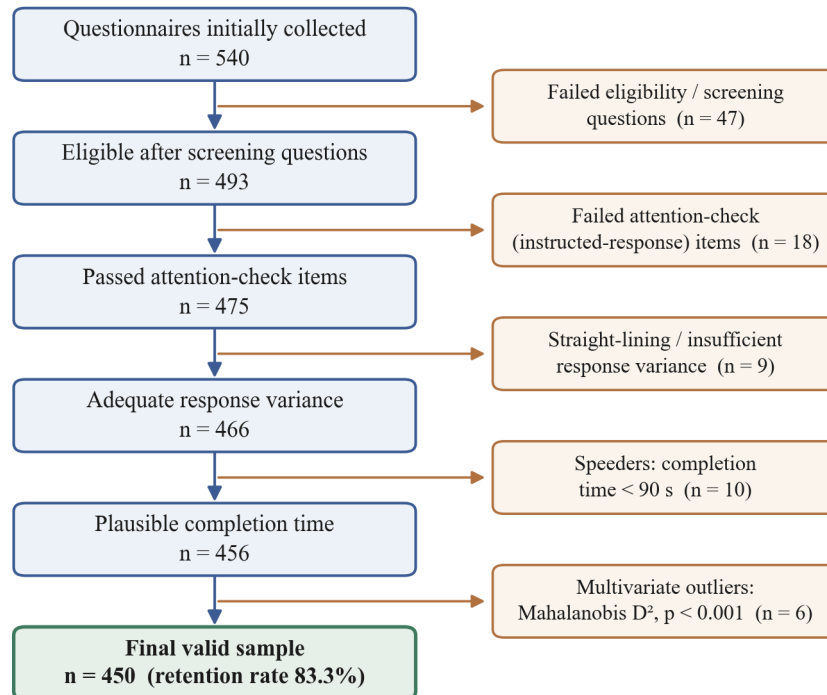


Fig.2: Data-screening cascade from the initially collected questionnaires to the final valid sample (N = 450).

Item-level descriptive statistics are reported in Table 4. All item means fell in the medium-to-high range (4.46–4.62 on the seven-point scale), consistent with the practical tendency for respondents to evaluate a recent successful interaction relatively positively. Skewness ranged from -0.37 to -0.18 and

kurtosis from -0.46 to -0.09 ; all values lay well within the ± 1 range, indicating no severe departure from normality. As PLS-SEM is a nonparametric, variance-based method, it is in any case robust to moderate non-normality. The standardised outer loadings and outer VIF values reported in Table 4 are discussed together with the measurement-model assessment in Section 4.3. Figure 3 profiles the item means with ± 1 SD intervals alongside the standardised outer loadings.

Table 4. Item-Level Descriptive Statistics, Normality, Standardised Loadings, and Outer VIF

Construct	Item	<i>M</i>	<i>SD</i>	Skewness	Kurtosis	Loading (λ)	VIF
AI customer service experience	AIX1	4.61	0.95	−0.32	−0.18	0.771	1.78
	AIX2	4.58	0.98	−0.28	−0.35	0.736	1.62
	AIX3	4.49	1.02	−0.21	−0.41	0.754	1.71
	AIX4	4.52	0.96	−0.35	−0.12	0.748	1.66
	AIX5	4.50	1.01	−0.24	−0.39	0.726	1.55
Perceived diagnosticity	DIA1	4.55	1.00	−0.30	−0.22	0.740	1.58
	DIA2	4.50	0.99	−0.27	−0.31	0.785	1.83
	DIA3	4.57	1.03	−0.19	−0.44	0.718	1.49
	DIA4	4.50	0.98	−0.33	−0.15	0.754	1.66
Trust	TRU1	4.52	1.03	−0.26	−0.37	0.775	1.84
	TRU2	4.49	1.00	−0.31	−0.20	0.793	1.95
	TRU3	4.53	1.04	−0.22	−0.45	0.780	1.88
	TRU4	4.46	1.02	−0.18	−0.40	0.755	1.71
Satisfaction	SAT1	4.58	1.01	−0.34	−0.16	0.795	1.97
	SAT2	4.55	1.04	−0.25	−0.42	0.759	1.74
	SAT3	4.53	1.03	−0.28	−0.33	0.759	1.73
	SAT4	4.58	1.00	−0.37	−0.09	0.806	2.04
Brand loyalty	LOY1	4.62	1.02	−0.33	−0.19	0.763	1.73
	LOY2	4.57	1.05	−0.24	−0.46	0.754	1.68
	LOY3	4.60	1.01	−0.36	−0.13	0.751	1.66
	LOY4	4.57	1.04	−0.27	−0.38	0.777	1.83

Note. *M* = mean; *SD* = standard deviation; λ = standardised outer loading; VIF = variance inflation factor. All loadings exceed 0.70 and all outer VIF values are below the 3.3 threshold, indicating satisfactory indicator reliability and no indicator-level collinearity.

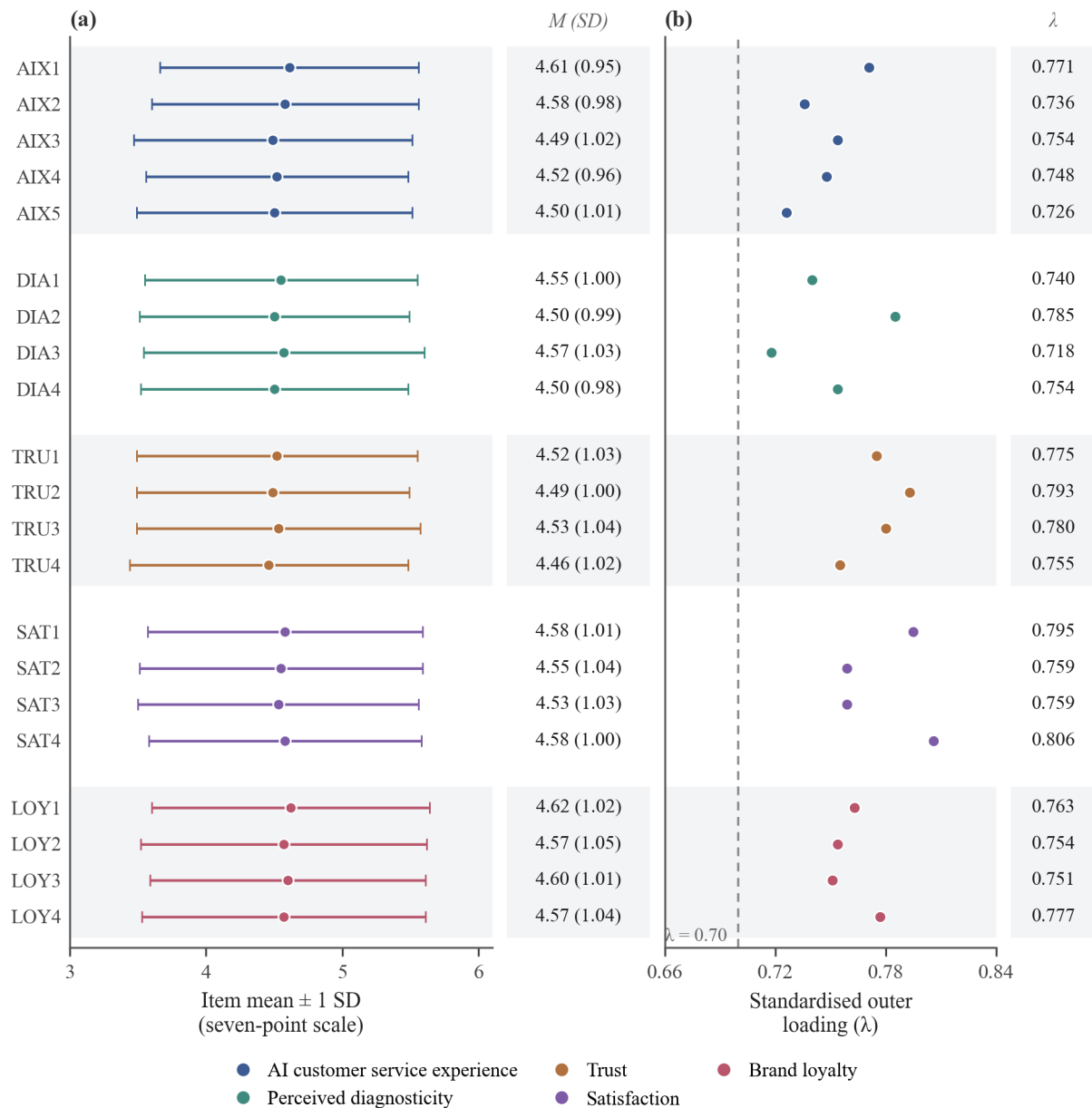


Fig.3: Item-level measurement profile: (a) item means with ± 1 SD intervals on the seven-point scale; (b) standardised outer loadings relative to the 0.70 threshold.

4.2. Common Method Bias

Because all measures were obtained from a single source, common method bias (CMB) was examined statistically. Harman's single-factor test showed that the first unrotated factor accounted for 34.27% of the total variance, which is well below the 50% threshold, indicating that no single factor dominated the covariance among the items. In addition, a full-collinearity assessment was conducted following Kock (2015): the variance inflation factors obtained by regressing each latent variable on all others ranged from 1.21 to 2.18, all below the conservative 3.3 criterion. Taken together with the procedural remedies described in Section 3.3, these results suggest that common method bias is unlikely to be a serious concern in this study.

4.3. Measurement Model Assessment

4.3.1. Reliability and Convergent Validity

The reliability and convergent validity of the reflective constructs are reported in Table 5. All

standardised outer loadings (Table 4) exceeded 0.70, ranging from 0.718 to 0.806, indicating adequate indicator reliability; the satisfaction construct displayed the highest loadings, reflecting high consistency in respondents' satisfaction assessments after an AI customer service interaction. At the construct level, Cronbach's α values ranged from 0.821 to 0.848, the ρ_A coefficients from 0.829 to 0.854, and the composite reliability (CR, ρ_c) values from 0.837 to 0.863. For every construct the estimates satisfied the expected ordering $\alpha \leq \rho_A \leq \rho_c$, and all exceeded the 0.70 benchmark, demonstrating strong internal-consistency reliability. The average variance extracted (AVE) ranged from 0.558 to 0.609, exceeding the 0.50 threshold and thereby confirming convergent validity for every construct. Figure 4 visualises these reliability and convergent-validity indices against their recommended thresholds.

Table 5. Construct Reliability and Convergent Validity

Construct	<i>M</i>	<i>SD</i>	Cronbach's α	ρ_A	CR	AVE
AI customer service experience	4.54	0.93	0.848	0.852	0.863	0.558
Perceived diagnosticity	4.53	0.97	0.821	0.829	0.837	0.562
Trust	4.50	1.01	0.844	0.850	0.858	0.602
Satisfaction	4.56	1.02	0.847	0.854	0.861	0.609
Brand loyalty	4.59	1.03	0.831	0.839	0.846	0.580

Note. *M* = mean; *SD* = standard deviation; ρ_A = Dijkstra–Henseler reliability coefficient; CR = composite reliability; AVE = average variance extracted. All α , ρ_A and CR values exceed 0.70 and all AVE values exceed 0.50.

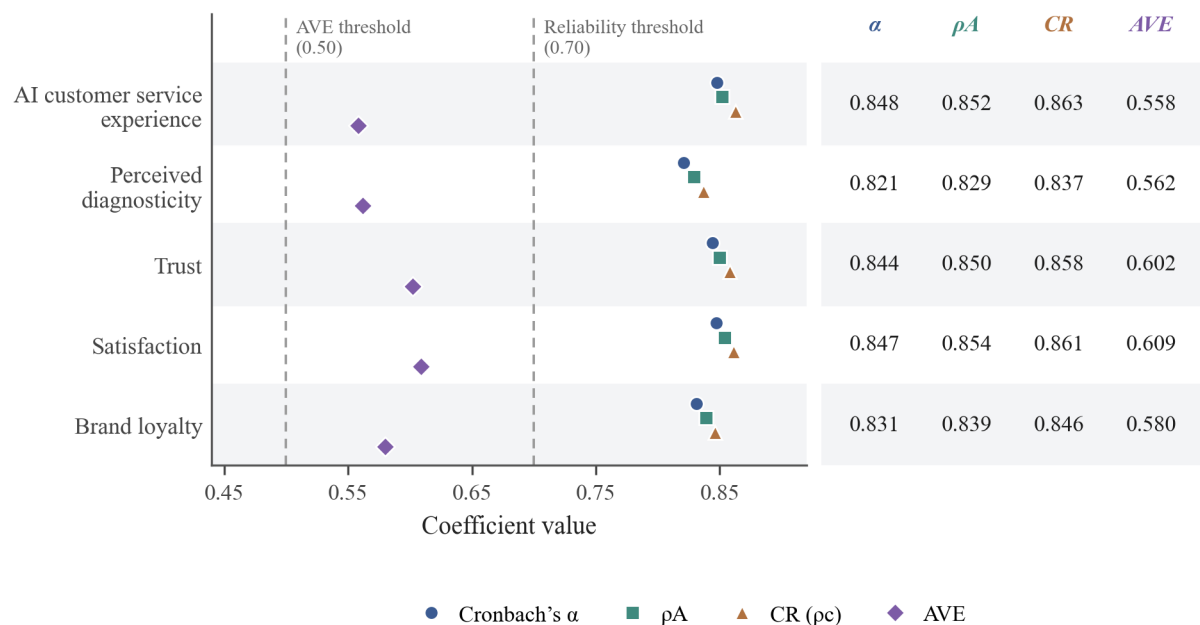


Fig.4: Construct reliability and convergent validity relative to the recommended thresholds (0.70 for Cronbach's α , ρ_A and CR; 0.50 for AVE).

4.3.2. Discriminant Validity

Discriminant validity was assessed using the Fornell–Larcker criterion and the HTMT ratio. As shown in Table 6, the square root of the AVE for each construct (on the diagonal) exceeded that construct's correlations with all other constructs, satisfying the Fornell–Larcker criterion. The inter-construct correlations were all positive and theoretically consistent: the correlation between AI customer service

experience and perceived diagnosticity was 0.651, and that between perceived diagnosticity and trust was 0.645, whereas the correlations among trust, satisfaction and brand loyalty ranged from 0.455 to 0.551. All correlations were below 0.80, indicating no obvious construct overlap. As reported in Table 7, every HTMT value was below the conservative 0.85 threshold, with the highest value being 0.723 (AI customer service experience–perceived diagnosticity). The two criteria jointly confirm that the five constructs are empirically distinct. Figure 5 integrates the two criteria in a single matrix display.

Table 6. Discriminant Validity: Fornell–Larcker Criterion

Construct	AICSE	PD	TR	SA	BL
AICSE	0.747				
PD	0.651	0.750			
TR	0.494	0.645	0.776		
SA	0.483	0.423	0.507	0.780	
BL	0.415	0.410	0.455	0.551	0.762

Note. Diagonal values (in bold) are the square root of the AVE; off-diagonal values are inter-construct correlations. AICSE = AI customer service experience; PD = perceived diagnosticity; TR = trust; SA = satisfaction; BL = brand loyalty.

Table 7. Discriminant Validity: Heterotrait–Monotrait Ratio (HTMT)

Construct	AICSE	PD	TR	SA	BL
AICSE	n/a				
PD	0.723	n/a			
TR	0.548	0.716	n/a		
SA	0.536	0.470	0.563	n/a	
BL	0.461	0.455	0.505	0.612	n/a

Note. All HTMT values are below the conservative threshold of 0.85. AICSE = AI customer service experience; PD = perceived diagnosticity; TR = trust; SA = satisfaction; BL = brand loyalty.

	AICSE	PD	TR	SA	BL
AICSE	0.747	0.723	0.548	0.536	0.461
PD	0.651	0.750	0.716	0.470	0.455
TR	0.494	0.645	0.776	0.563	0.505
SA	0.483	0.423	0.507	0.780	0.612
BL	0.415	0.410	0.455	0.551	0.762

Fig.5: Combined discriminant-validity matrix: $\sqrt{\text{AVE}}$ on the diagonal, inter-construct correlations below the diagonal (Fornell–Larcker criterion), and HTMT ratios above the diagonal.

Note. AICSE = AI customer service experience; PD = perceived diagnosticity; TR = trust; SA = satisfaction; BL = brand loyalty. Bold diagonal cells report the $\sqrt{\text{AVE}}$ of each construct. All HTMT values lie below the conservative 0.85 threshold.

4.4. Structural Model Assessment

4.4.1. Collinearity Assessment

Before interpreting the path coefficients, the inner-model collinearity was examined. As shown in Table 8, all inner VIF values ranged from 1.000 to 1.511, far below the 3.3 threshold. Collinearity among the predictor constructs is therefore not a concern, and the structural-path estimates can be interpreted without bias.

Table 8. Structural-Model Collinearity (Inner VIF)

Predictor construct	Outcome construct	VIF
AI customer service experience	Perceived diagnosticity	1.000
Perceived diagnosticity	Trust	1.000
AI customer service experience	Satisfaction	1.323
Trust	Satisfaction	1.323
AI customer service experience	Brand loyalty	1.464
Trust	Brand loyalty	1.511
Satisfaction	Brand loyalty	1.489

Note. VIF = variance inflation factor. All values are below the 3.3 threshold, indicating no critical collinearity in the structural model.

4.4.2. Path Coefficients and Hypothesis Testing

The standardised path coefficients and the associated bootstrap statistics are reported in Table 9 and visualised in Figure 6. All seven direct paths were positive and statistically significant, and thus all hypotheses were supported. AI customer service experience was strongly associated with perceived diagnosticity ($\beta = 0.651$, $t = 21.00$, $p < 0.001$; H2), indicating that the more positively consumers evaluate the information clarity, response timeliness and problem-solving capability of AI customer service, the higher their perceived diagnosticity regarding cosmetics fit, product differences and purchase judgement. Perceived diagnosticity in turn had a positive effect on trust ($\beta = 0.561$, $t = 14.76$, $p < 0.001$; H3), showing that when AI customer service helps consumers reduce uncertainty in product selection, they are more likely to trust the brand and its service information. Satisfaction was jointly driven by AI customer service experience ($\beta = 0.308$, $t = 6.55$, $p < 0.001$; H4) and trust ($\beta = 0.354$, $t = 7.70$, $p < 0.001$; H5). With respect to brand loyalty, AI customer service experience ($\beta = 0.131$, $t = 2.52$, $p = 0.012$; H1), trust ($\beta = 0.193$, $t = 3.78$, $p < 0.001$; H6) and satisfaction ($\beta = 0.390$, $t = 8.13$, $p < 0.001$; H7) all had positive effects, with satisfaction exerting the strongest influence. These results indicate that AI customer service experience directly affects brand loyalty and also indirectly shapes the brand relationship by enhancing perceived diagnosticity, trust and satisfaction. Figure 7 complements the model diagram by displaying the bias-corrected 95% bootstrap confidence interval for each structural path.

Table 9. Structural-Path Estimates and Hypothesis Testing

H	Path	β	t	p	95% CI	F^2	Decision
H1	AICSE $\rightarrow$ Brand loyalty	0.131	2.52	0.012	[0.028, 0.232]	0.019	Supported
H2	AICSE $\rightarrow$ Perceived diagnosticity	0.651	21.00	<0.001	[0.589, 0.710]	0.736	Supported
H3	Perceived diagnosticity $\rightarrow$ Trust	0.561	14.76	<0.001	[0.484, 0.633]	0.460	Supported
H4	AICSE $\rightarrow$ Satisfaction	0.308	6.55	<0.001	[0.215, 0.398]	0.106	Supported
H5	Trust $\rightarrow$ Satisfaction	0.354	7.70	<0.001	[0.262, 0.442]	0.141	Supported
H6	Trust $\rightarrow$ Brand loyalty	0.193	3.78	<0.001	[0.094, 0.293]	0.038	Supported
H7	Satisfaction $\rightarrow$ Brand loyalty	0.390	8.13	<0.001	[0.296, 0.483]	0.159	Supported

Note. AICSE = AI customer service experience. β = standardised path coefficient; F^2 = effect size (0.02 = small, 0.15 = medium, 0.35 = large). Bias-corrected 95% confidence intervals (CI) are based on 5,000 bootstrap subsamples; none includes zero.

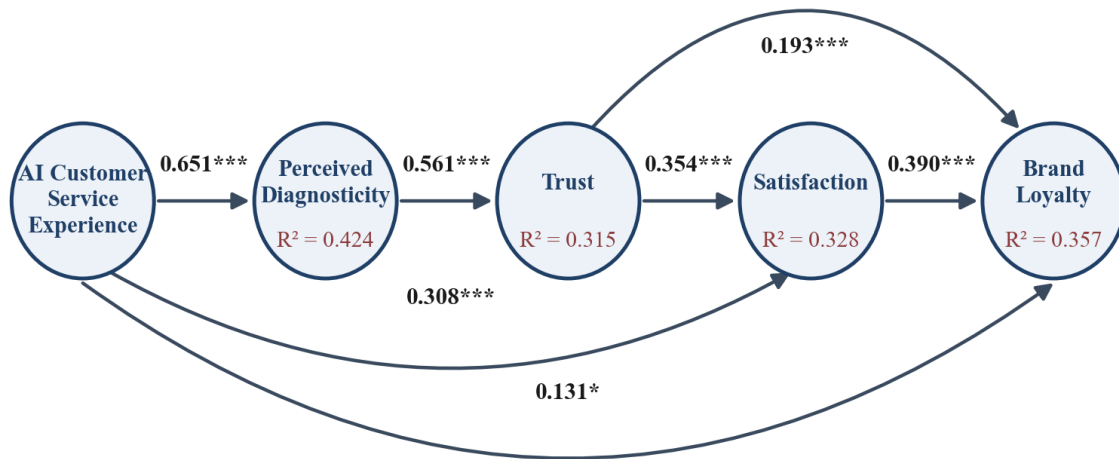


Fig.6: Final structural model with standardised path coefficients (PLS-SEM results).

Note. Values on paths are standardised path coefficients (β); R^2 is reported inside the endogenous constructs. Significance (two-tailed, bias-corrected bootstrap with 5,000 subsamples): * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

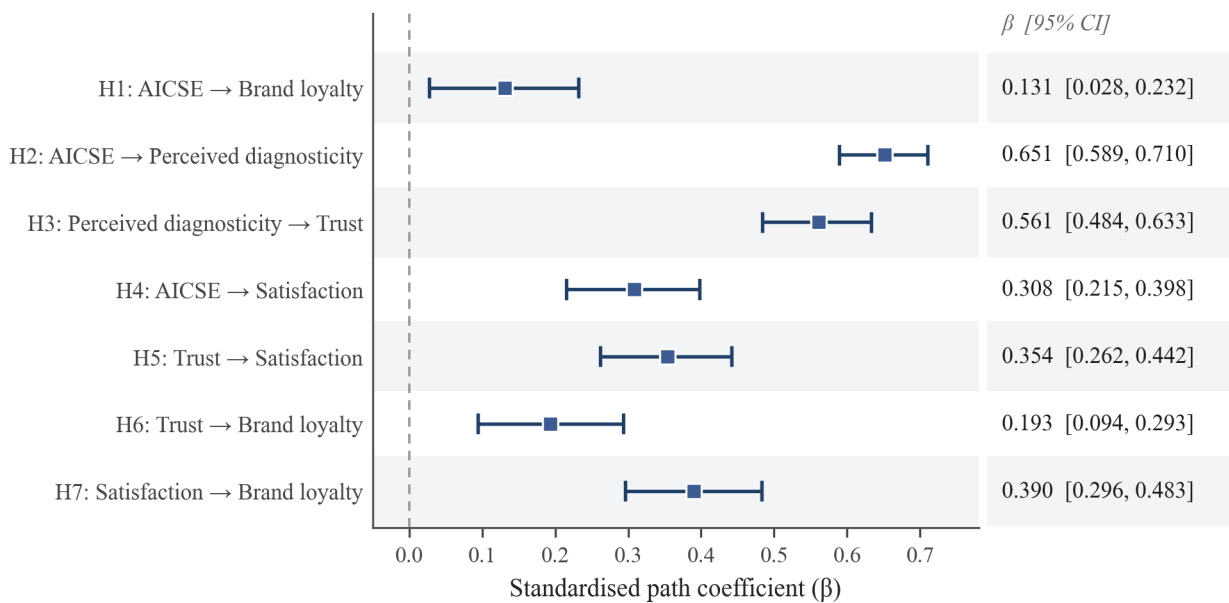


Fig.7: Standardised structural-path coefficients with bias-corrected 95% bootstrap confidence intervals (5,000 subsamples).

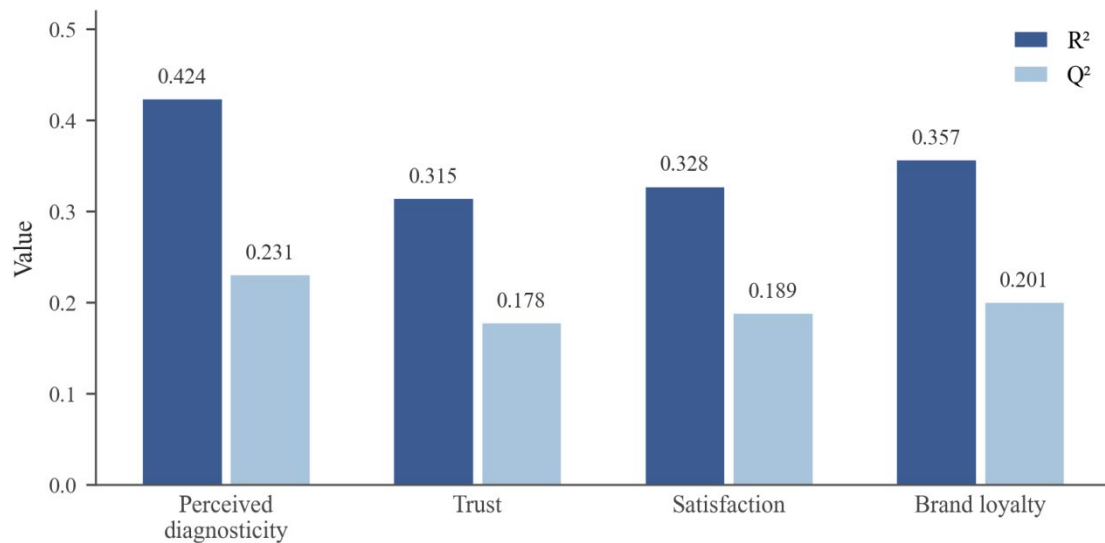
4.4.3. Explanatory Power and Predictive Relevance

The model accounted for a substantial share of variance in the endogenous constructs (Table 10). The R^2 values were 0.424 for perceived diagnosticity, 0.315 for trust, 0.328 for satisfaction, and 0.357 for brand loyalty, indicating moderate explanatory power. The effect sizes (f^2 in Table 9) confirmed that AI customer service experience had a large effect on perceived diagnosticity ($f^2 = 0.736$) and that perceived diagnosticity had a large effect on trust ($f^2 = 0.460$), whereas satisfaction had a medium-to-large effect on brand loyalty ($f^2 = 0.159$). Predictive relevance was assessed through blindfolding: all Stone–Geisser Q^2 values were clearly above zero (0.178–0.231), confirming the model's out-of-sample predictive relevance for every endogenous construct. Figure 8 juxtaposes the R^2 and Q^2 values across the four endogenous constructs.

Table 10. Coefficient of Determination (R^2) and Predictive Relevance (Q^2)

Endogenous construct	R^2	R^2 adjusted	Q^2
Perceived diagnosticity	0.424	0.423	0.231
Trust	0.315	0.313	0.178
Satisfaction	0.328	0.325	0.189
Brand loyalty	0.357	0.353	0.201

Note. R^2 = coefficient of determination; Q^2 = Stone–Geisser predictive relevance obtained by blindfolding (omission distance = 7). $Q^2 > 0$ indicates predictive relevance.

Fig.8: Explanatory power (R^2) and predictive relevance (Q^2) of the endogenous constructs.

4.4.4. Model Fit

The approximate model-fit indices are reported in Table 11. The standardised root mean square residual (SRMR) was 0.045 for the saturated model and 0.052 for the estimated model, both below the 0.08 threshold. The normed fit index (NFI) was 0.913 and 0.907, respectively, above the 0.90 benchmark, and the discrepancy measures d_{ULS} and d_G were within their respective bootstrap-based confidence limits. In addition, the root mean square residual covariance (rms_theta) was 0.118, below the 0.12 criterion. Collectively, these indices indicate an acceptable overall model fit.

These approximate indices are reported for completeness. Because PLS-SEM is a variance-based, prediction-oriented method, model evaluation rests primarily on the reliability and validity of the measurement model and on the structural model's explanatory power (R^2), effect sizes (f^2), and out-of-sample predictive relevance (Q^2) rather than on global goodness-of-fit indices, whose status in PLS-SEM remains debated (Hair et al., 2019). The fit statistics above should therefore be read as supplementary rather than primary evidence.

Table 11. Model-Fit Indices

Index	Saturated model	Estimated model	Recommended
SRMR	0.045	0.052	< 0.08
d_{ULS}	0.486	0.654	within CI
d_G	0.187	0.205	within CI

Chi-square (χ^2)	612.44	657.22	smaller is better
NFI	0.913	0.907	> 0.90

Note. SRMR = standardised root mean square residual; d_ULS = squared Euclidean distance; d_G = geodesic distance; NFI = normed fit index. The model-level rms_theta was 0.118 (< 0.12).

4.4.5. Mediation Analysis

The mediation effects were tested with 5,000 bootstrap resamples; the bias-corrected 95% confidence intervals for all indirect effects excluded zero, confirming that the mediating paths were statistically significant (Table 12). The indirect effect of “AI customer service experience → satisfaction → brand loyalty” was the largest specific indirect effect ($\beta = 0.120$, 95% CI [0.078, 0.163]), indicating that consumers’ satisfaction with the AI customer service interaction is a crucial mechanism transferring service experience into brand loyalty. The indirect effect of “AI customer service experience → perceived diagnosticity → trust → brand loyalty” was also significant ($\beta = 0.070$, 95% CI [0.036, 0.107]), showing that AI customer service fosters brand trust by enhancing perceived diagnosticity, which in turn cultivates loyalty. The full serial path “AI customer service experience → perceived diagnosticity → trust → satisfaction → brand loyalty” was significant as well ($\beta = 0.051$, 95% CI [0.033, 0.070]), which is consistent with serial mediation (H8). The total indirect effect was 0.241 and the total effect was 0.372 (95% CI [0.284, 0.453]). Because the direct effect of AI customer service experience on brand loyalty ($\beta = 0.131$) and the indirect effects were both significant and of the same sign, the results indicate complementary (partial) mediation. In sum, the influence of AI customer service experience on brand loyalty is not a single psychological response but a continuous transmission process of enhanced decision support, brand-trust formation, satisfaction improvement, and loyalty accumulation. Figure 9 visualizes this decomposition of the total effect into its direct and indirect components.

For transparency, all model-implied indirect effects were estimated; Table 12 reports the theoretically focal specific indirect effects together with the total indirect and total effects, while the remaining lower-order specific indirect effects were consistent in sign and significance and are available on request. Reporting the focal paths avoids redundancy while still documenting the full decomposition of the total effect. Because the design is cross-sectional, these results should be interpreted as consistent with the hypothesised serial ordering rather than as establishing the temporal sequence among the mediators; the analysis identifies a theoretically ordered set of associations but cannot, on its own, confirm their causal direction.

Table 12. Direct, Indirect, and Total Effects (Bootstrap, 5,000 Subsamples)

Effect	β	t	p	95% CI	Conclusion
Direct: AICSE → BL (H1)	0.131	2.52	0.012	[0.028, 0.232]	Significant
Indirect: AICSE → SA → BL	0.120	5.53	<0.001	[0.078, 0.163]	Supported
Indirect: AICSE → PD → TR → BL	0.070	3.87	<0.001	[0.036, 0.107]	Supported
Serial: AICSE → PD → TR → SA → BL (H8)	0.051	5.40	<0.001	[0.033, 0.070]	Supported
Total indirect effect	0.241	7.10	<0.001	[0.176, 0.309]	n/a
Total effect: AICSE → BL	0.372	8.63	<0.001	[0.284, 0.453]	n/a

Note. AICSE = AI customer service experience; PD = perceived diagnosticity; TR = trust; SA = satisfaction; BL = brand loyalty. Bias-corrected 95% confidence intervals are based on 5,000 bootstrap subsamples.

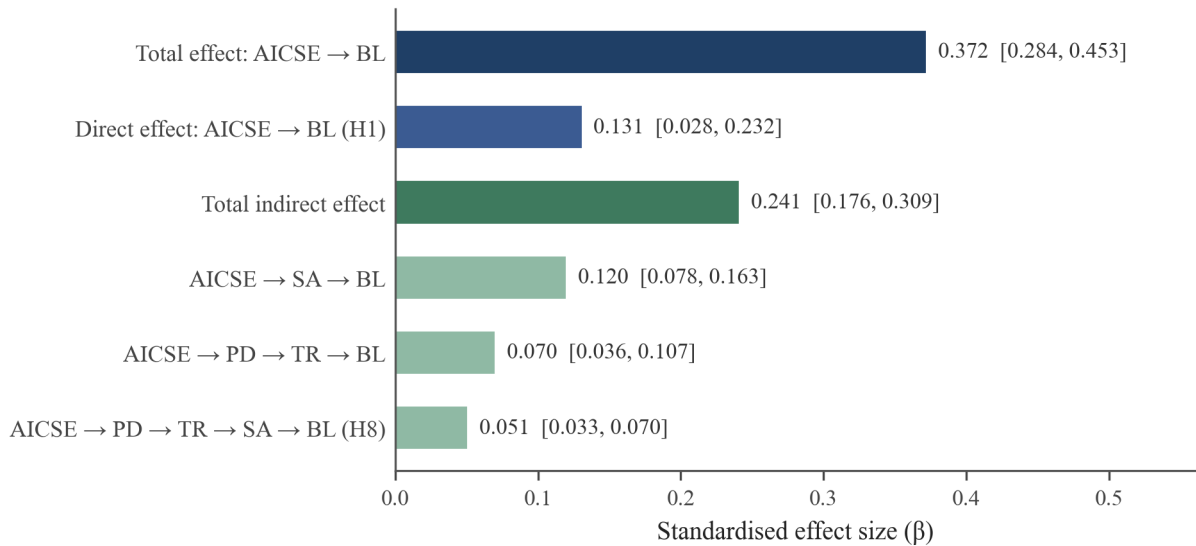


Fig.9: Decomposition of the total effect of AI customer service experience on brand loyalty into its direct and indirect components (bias-corrected 95% bootstrap confidence intervals in brackets).

4.4.6. Robustness Checks: Control Variables and Prior Brand Experience

To assess robustness, the structural model was re-estimated with seven control variables (age, gender, education, monthly disposable income, most-recent purchase category, live-streaming platform, and prior brand-use experience) specified as additional predictors of brand loyalty. The control-variable results are reported in Table 13, and the coefficient-stability comparison is reported in Table 14. The focal structural paths retained their direction and statistical significance, and their coefficients changed only marginally (all $|\Delta\beta| < 0.017$), indicating that the hypothesised relationships are not artefacts of sample composition. Among the controls, only prior brand-use experience was a non-trivial predictor of brand loyalty ($\beta = 0.087$, $p = 0.028$); the remaining controls were not statistically significant.

Table 13. Control-Variable Model Predicting Brand Loyalty

Predictor	β	t	p	VIF	Interpretation
AI customer service experience	0.126	2.44	0.015	1.72	Significant
Trust	0.184	3.56	<0.001	1.42	Significant
Satisfaction	0.374	7.72	<0.001	1.38	Significant
Prior brand-use experience	0.087	2.20	0.028	1.19	Significant control
Age	-0.018	0.47	0.642	1.12	Not significant
Gender	0.021	0.55	0.584	1.08	Not significant
Education	0.029	0.77	0.441	1.31	Not significant
Monthly disposable income	0.034	0.92	0.358	1.35	Not significant
Most-recent purchase category	0.025	0.67	0.506	1.18	Not significant
Live-streaming platform	-0.011	0.29	0.773	1.16	Not significant

Note. Brand loyalty R^2 increased from 0.357 to 0.372 after adding the controls. Prior brand-use experience was coded 1 = previous experience with the focal brand and 0 = no previous experience.

Table 14. Stability of Focal Brand-Loyalty Paths after Adding Controls

Path	Baseline β	With controls β	$\Delta\beta$
AI customer service experience $\rightarrow$ Brand loyalty	0.131	0.126	-0.005
Trust $\rightarrow$ Brand loyalty	0.193	0.184	-0.009
Satisfaction $\rightarrow$ Brand loyalty	0.390	0.374	-0.016

Note. The largest absolute change in a focal coefficient was 0.016, supporting coefficient stability after the inclusion of demographic, purchase-category, platform and prior-experience controls.

Because prior brand-use experience may itself drive loyalty, a partial least squares multigroup analysis (PLS-MGA) compared consumers with and without prior experience of the focal brand ($n = 303$ vs. $n = 147$). Measurement invariance was supported by the MICOM procedure (configural and compositional invariance; Table 15), and none of the focal path differences across the two groups was statistically significant (all PLS-MGA p values ≥ 0.124 ; Table 16), indicating that the proposed mechanism operates consistently regardless of prior brand experience. Figure 10 contrasts the group-specific path coefficients, whose profiles are closely aligned across the two groups.

Table 15. MICOM Compositional Invariance by Prior Brand Experience

Construct	Step 1	c value	5% quantile	Permutation p	Conclusion
AICSE	Established	0.999	0.994	0.382	Compositional invariance
Perceived diagnosticity	Established	0.998	0.993	0.417	Compositional invariance
Trust	Established	0.999	0.995	0.364	Compositional invariance
Satisfaction	Established	0.998	0.992	0.446	Compositional invariance
Brand loyalty	Established	0.999	0.994	0.401	Compositional invariance

Note. AICSE = AI customer service experience. In MICOM step 2, $p > 0.05$ indicates that compositional invariance is not rejected.

Table 16. PLS-MGA Path Differences by Prior Brand Experience

Path	Prior experience $n = 303$	No prior experience $n = 147$	Difference	PLS-MGA p
AICSE $\rightarrow$ Brand loyalty	0.124	0.149	-0.025	0.624
AICSE $\rightarrow$ Perceived diagnosticity	0.663	0.629	0.034	0.317
Perceived diagnosticity $\rightarrow$ Trust	0.548	0.584	-0.036	0.408
AICSE $\rightarrow$ Satisfaction	0.294	0.333	-0.039	0.382
Trust $\rightarrow$ Satisfaction	0.365	0.329	0.036	0.437
Trust $\rightarrow$ Brand loyalty	0.205	0.171	0.034	0.511
Satisfaction $\rightarrow$ Brand loyalty	0.373	0.434	-0.061	0.124

Note. Two-tailed PLS-MGA p values are reported. None of the focal group differences reached $p < 0.05$.

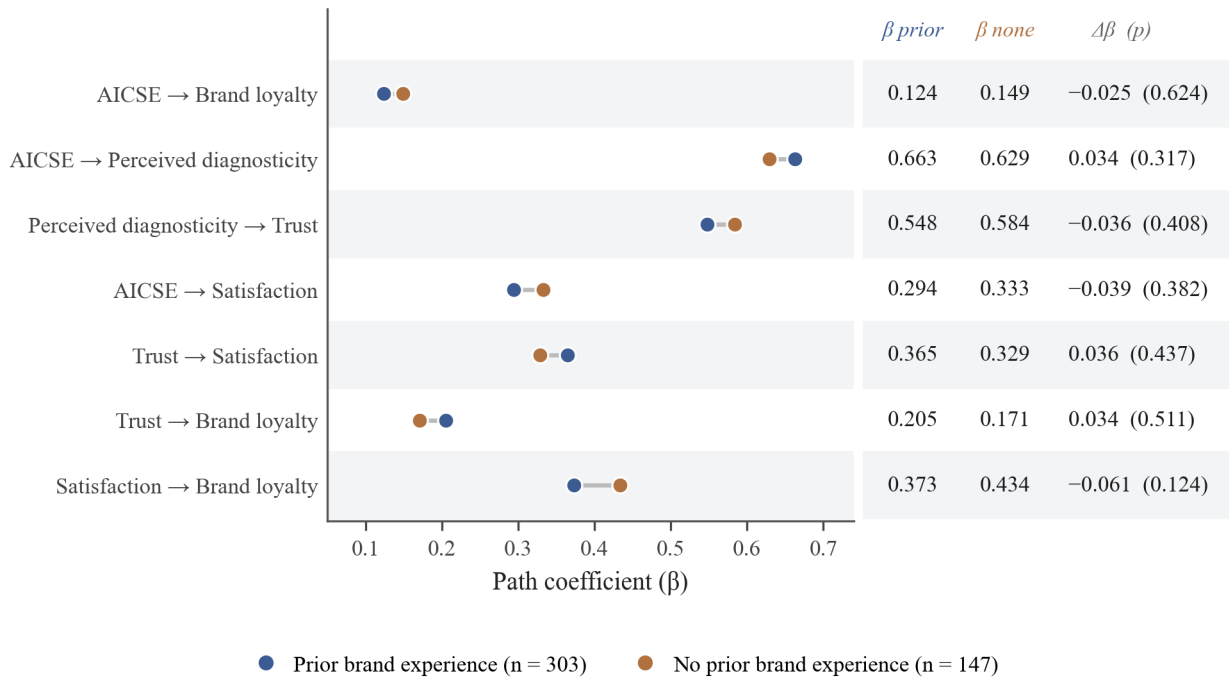


Fig.10: PLS-MGA comparison of structural paths between consumers with (n = 303) and without (n = 147) prior brand experience.

5. Discussion

5.1. Main Findings

The study shows that AI customer service experience positively influences consumer brand loyalty. In the Chinese cosmetics live-streaming commerce context, AI customer service is not only an auxiliary tool for instant consultation but also an important digital touchpoint that shapes consumers' brand-relationship judgements. When consumers find AI customer service to be clear, timely, effective and relevant to their needs, they are more likely to continue to choose, repurchase from, and recommend the brand.

The findings also reveal that perceived diagnosticity is a critical cognitive mechanism explaining the effect of AI customer service experience on brand loyalty. The value of AI customer service lies not only in the speed and convenience of service for cosmetics consumers, but, more importantly, in whether it helps them judge how well a product suits their skin type, functional needs, shade preference and usage scenario. In other words, high-quality AI customer service can reduce product-fit uncertainty and provide more solid reasons for purchase judgement.

The study further shows that trust and satisfaction form a sequential transmission relationship in the influence of AI customer service experience on brand loyalty. Consumers are more likely to develop trust in the brand and its service information when they perceive the diagnosticity of information provided by AI customer service to be high. Such trust further enhances their satisfaction with the interaction process and shopping experience, which is finally manifested in loyalty outcomes such as repurchase, recommendation, and continued preference. In sum, the impact of AI customer service on brand loyalty is not limited to the surface level of service experience but gradually develops into a long-term brand relationship through diagnostic judgement, trust formation and satisfaction evaluation.

Two qualifications sharpen this interpretation. First, the direct association between AI customer service experience and brand loyalty is comparatively small ($\beta = 0.131$; $f^2 = 0.019$); the total effect operates largely through perceived diagnosticity, trust, and satisfaction, so that AI customer service appears to build loyalty mainly by improving decision support and the relational states it fosters rather

than through a direct route. Second, because the data are cross-sectional, these pathways should be read as associational rather than strictly causal: the analysis describes a theoretically ordered set of relationships that is consistent with, but cannot by itself establish, the proposed temporal sequence.

5.2. Theoretical Contributions

The first contribution of this study is that it expands AI customer service research from general e-commerce to the more contextually sensitive area of cosmetics live streaming. Previous research has demonstrated that live streaming can reduce uncertainty in apparel and cosmetics consumption and that interaction, social presence and dialogic quality can affect trust and behavioural responses; however, relatively little research has examined the operation of AI customer service inside live-streaming rooms (Lu & Chen, 2021; Jiang et al., 2022; Konya-Baumbach et al., 2023). This study treats AI customer service as an independent digital service touchpoint and complements the streamer-centred perspective in live-streaming research.

Second, this study incorporates perceived diagnosticity into the pathway through which AI customer service influences loyalty, filling the gap of insufficient attention to decision-support mechanisms for experiential products in the traditional “service quality → satisfaction → loyalty” framework. In the cosmetics context, consumers face not merely an information gap but a product-fit judgement problem. Perceived diagnosticity therefore explains the effectiveness of AI customer service better than information richness does.

Third, this study reconfirms the importance of distinguishing between continued use of a technology and long-term loyalty toward a brand. Whereas some prior studies of AI chatbots have focused on continued intention to use the chatbot itself, this study takes brand loyalty as the dependent variable, bringing the analysis closer to relationship marketing and long-term firm value.

Read against prior work, these results both converge with and extend earlier findings. Consistent with research showing that AI service quality raises loyalty through perceived value, trust, and satisfaction in general e-commerce (Chen et al., 2023; Hsu & Lin, 2023; Huang et al., 2024), we observe similar relational mechanisms in cosmetics live commerce; however, the comparatively small direct path and the dominant indirect transmission through perceived diagnosticity indicate that, in this high-uncertainty setting, decision support, rather than service quality in general, is the key mechanism. This refines the largely streamer-centred account of live commerce (Lu & Chen, 2021; Gao et al., 2023; Zhang et al., 2026) by identifying AI customer service as an additional, decision-oriented service touchpoint within the platform’s service system.

5.3. Managerial Implications

Brands should not treat AI customer service primarily as a cost-cutting tool. Instead, it should be designed as a decision-support system inside live-streaming rooms. For cosmetics, it is more helpful to match skin types, recommend shades, and explain ingredients, usage scenarios and combination advice than to offer generic greetings.

Second, if brands want to improve loyalty through AI customer service, they need to prioritise information quality, problem-solving capability and personalised advice, rather than focusing narrowly on making AI “as human-like as possible”. Recent AI customer service research also suggests that information quality and problem-solving ability are more reliable drivers of satisfaction than superficial anthropomorphic cues (Hoang & Tran, 2026). The effect of emotional expression, humour and informal language in high-emotion and service-recovery contexts may, however, depend on boundary conditions such as service-failure severity (Rita et al., 2026; Zhang et al., 2024; Shams et al., 2024).

Third, AI customer service should be consistent with the promises made by the streamer, the product detail pages, and after-sales service. When live explanations, customer-service responses and actual service delivery contradict one another, consumer trust in the brand can easily fall back to the

starting point or even deteriorate.

More concretely, brands can operationalise AI customer service as a decision-support workflow: a structured intake of skin type and skin concerns; shade- and product-matching prompts that translate stated needs into specific recommendations; ingredient and suitability checks that flag potential irritants or contraindications; and clear escalation to human agents when a query exceeds the system's competence or involves dissatisfaction and service recovery. From a service-operations standpoint, the AI layer should be integrated with inventory, promotion, and after-sales systems and governed by consistency checks that align streamer claims, product-detail pages, and AI responses, because contradictions across these touchpoints are precisely where trust erodes. This reframes AI customer service from a cost-reduction channel into a managed component of the platform's digital service system.

6. Limitations and Future Research

Despite using real questionnaire data to explore the mechanism through which AI customer service experience influences consumer brand loyalty in cosmetics live-streaming commerce, this study has several limitations. First, cross-sectional questionnaire data were used. Such data can reveal correlations and theoretical paths among variables but provide little insight into the dynamic evolution among AI customer service experience, trust, satisfaction, and brand loyalty. Repeated live-streaming shopping experiences, brand-fulfilment performance and after-sales service may all influence consumers' evaluation of AI customer service. Future research could employ longitudinal tracking designs or experimental approaches to examine trust building, satisfaction change, and loyalty development at different time points.

Second, the study mainly focused on pre-sale and in-sale consultation in cosmetics live-streaming shopping and did not sufficiently incorporate after-sales situations such as complaints, returns and exchanges, service recovery, and negative-experience management. Previous studies have suggested that the role of chatbots in service failure, high-emotion and customer-dissatisfaction contexts may differ from their role in routine consultation. Future research could therefore compare routine consultation with service-recovery tasks to explore the differentiated effects of AI customer service on consumer trust, satisfaction, and brand loyalty under different service-pressure conditions.

Third, brand loyalty was used as an outcome variable focusing on consumers' continued brand preference, repurchase, and recommendation tendencies. Future research could deconstruct brand loyalty into multiple dimensions such as repurchase intention, recommendation intention, price-premium acceptance, brand tolerance, and cross-channel continued purchasing. Such differentiation would enable a more accurate identification of the effects of AI customer service experience on different loyalty outcomes. In cosmetics consumption, consumers may repurchase a particular brand without accepting a higher price, or may be willing to recommend only a single product without showing overall brand loyalty. Thus, disaggregating brand loyalty could improve the explanatory power of future research.

Fourth, this study investigated the chain mechanism of AI customer service experience on brand loyalty through perceived diagnosticity, trust, and satisfaction, but did not examine boundary conditions such as consumer and product characteristics in depth. Future research could consider moderators such as perceived product uncertainty, product involvement, prior brand experience, AI technology acceptance, cosmetics knowledge, or privacy concerns. This would help to clarify which consumer groups, product-judgement tasks and live-interaction scenarios make AI customer service more likely to function effectively.

Fifth, several features of the design constrain generalisation and warrant caution. The sample was obtained through non-probability online recruitment, so it captures, rather than statistically represents, the relevant consumer segments. The measures are self-reports collected after the interaction and are

therefore subject to recall and common-method biases; although Harman's single-factor test and a full-collinearity assessment indicate that these biases are not severe, residual common-method variance cannot be ruled out entirely. Moreover, the outcomes are perceptual and intentional rather than behavioural, as the study does not observe actual repurchase or recommendation behaviour. Future work could combine probability or quota sampling, multi-source or temporally separated measurement, and behavioural or platform-log data to strengthen both causal and external validity.

References

- Chen, Q., Lu, Y., Gong, Y., & Xiong, J. (2023). Can AI chatbots help retain customers? Impact of AI service quality on customer loyalty. *Internet Research*, 33(6), 2205–2243. <https://doi.org/10.1108/INTR-09-2021-0686>
- Cui, X., Law, M., Ng, M., & Lam, L. (2024). What drives consumers to buy in live streaming commerce? A systematic literature review. *Journal of Logistics, Informatics and Service Science*, 11(2), 464–486. <https://doi.org/10.33168/JLISS.2024.0229>
- Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly*, 39(2), 297–316. <https://doi.org/10.25300/MISQ/2015/39.2.02>
- Eni, Y., Princes, E., Hidayat, D., Zulaykha, A., Gunawan, A., & Jason. (2024). Investigating e-service quality, ease of use, trust and satisfaction in Indonesian e-commerce platforms. *Journal of Logistics, Informatics and Service Science*, 11(5), 292–301. <https://doi.org/10.33168/JLISS.2024.0517>
- Gao, W., Jiang, N., & Guo, Q. (2023). How do virtual streamers affect purchase intention in the live streaming context? A presence perspective. *Journal of Retailing and Consumer Services*, 73, Article 103356. <https://doi.org/10.1016/j.jretconser.2023.103356>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hajli, N., Sims, J., Zadeh, A. H., & Richard, M.-O. (2017). A social commerce investigation of the role of trust in a social networking site on purchase intentions. *Journal of Business Research*, 71, 133–141. <https://doi.org/10.1016/j.jbusres.2016.10.004>
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Hoang, C. C., & Tran, N. T. (2026). Chatbot self-presentation strategies and customer service satisfaction in e-commerce: The dual mediating roles of perceived competence and perceived warmth. *Journal of Logistics, Informatics and Service Science*, 13(5), 83–96. <https://doi.org/10.33168/JLISS.2026.0506>
- Hsu, C.-L., & Lin, J. C.-C. (2023). Understanding the user satisfaction and loyalty of customer service chatbots. *Journal of Retailing and Consumer Services*, 71, Article 103211. <https://doi.org/10.1016/j.jretconser.2022.103211>
- Huang, D., Markovitch, D. G., & Stough, R. A. (2024). Can chatbot customer service match human service agents on customer satisfaction? An investigation in the role of trust. *Journal of Retailing and Consumer Services*, 76, Article 103600. <https://doi.org/10.1016/j.jretconser.2023.103600>
- Jiang, H., Cheng, Y., Yang, J., & Gao, S. (2022). AI-powered chatbot communication with customers:

Dialogic interactions, satisfaction, engagement, and customer behavior. *Computers in Human Behavior*, 134, Article 107329. <https://doi.org/10.1016/j.chb.2022.107329>

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, 11(4), 1–10. <https://doi.org/10.4018/ijec.2015100101>

Konya-Baumbach, E., Biller, M., & von Janda, S. (2023). Someone out there? A study on the social presence of anthropomorphized chatbots. *Computers in Human Behavior*, 139, Article 107513. <https://doi.org/10.1016/j.chb.2022.107513>

Lu, B., & Chen, Z. (2021). Live streaming commerce and consumers' purchase intention: An uncertainty reduction perspective. *Information & Management*, 58(7), Article 103509. <https://doi.org/10.1016/j.im.2021.103509>

MacKenzie, S. B., & Podsakoff, P. M. (2012). Common method bias in marketing: Causes, mechanisms, and procedural remedies. *Journal of Retailing*, 88(4), 542–555. <https://doi.org/10.1016/j.jretai.2012.08.001>

Megdadi, O., Al-Ahmed, H., Ashour, M. L., Shriedeh, F. B., & Alshaketheep, K. (2025). The impact of chatbots on customer experience in e-commerce: Examining responsiveness, ease of use, and personalization. *Journal of Logistics, Informatics and Service Science*, 12(7), 147–163. <https://doi.org/10.33168/JLISS.2025.0709>

Meilatinova, N. (2021). Social commerce: Factors affecting customer repurchase and word-of-mouth intentions. *International Journal of Information Management*, 57, Article 102300. <https://doi.org/10.1016/j.ijinfomgt.2020.102300>

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing*, 58(3), 20–38. <https://doi.org/10.1177/002224299405800302>

Oliver, R. L. (1999). Whence consumer loyalty? *Journal of Marketing*, 63(4_suppl1), 33–44. <https://doi.org/10.1177/00222429990634s105>

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.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891. <https://doi.org/10.3758/BRM.40.3.879>

Rita, P., Vong, C., & Correia, C. (2026). How AI-driven chatbots shape customer satisfaction and loyalty to chatbot usage in digital service experience. *Computers in Human Behavior Reports*, 22, Article 101017. <https://doi.org/10.1016/j.chbr.2026.101017>

Shams, G., Kim, K. K., & Kim, K. (2024). Enhancing service recovery satisfaction with chatbots: The role of humor and informal language. *International Journal of Hospitality Management*, 120, Article 103782. <https://doi.org/10.1016/j.ijhm.2024.103782>

Wang, J.-C., & Chang, C.-H. (2013). How online social ties and product-related risks influence purchase intentions: A Facebook experiment. *Electronic Commerce Research and Applications*, 12(5), 337–346. <https://doi.org/10.1016/j.elerap.2013.03.003>

Zhang, J., Chen, Q., Lu, J., Wang, X., Liu, L., & Feng, Y. (2024). Emotional expression by artificial intelligence chatbots to improve customer satisfaction: Underlying mechanism and boundary conditions.

Tourism Management, 100, Article 104835. <https://doi.org/10.1016/j.tourman.2023.104835>

Zhang, R., Fang, J., Han, L., & Liao, M. (2026). Is more human-like better? The impact of virtual streamer language naturalness on user engagement. *Technological Forecasting and Social Change*, 222, Article 124392. <https://doi.org/10.1016/j.techfore.2025.124392>

Zhao, J.-D., Huang, J.-S., & Su, S. (2019). The effects of trust on consumers' continuous purchase intentions in C2C social commerce: A trust transfer perspective. *Journal of Retailing and Consumer Services*, 50, 42–49. <https://doi.org/10.1016/j.jretconser.2019.04.014>

Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>

Appendix

Table A1. Measurement Items

Construct	Code	Item	Source
AI customer service experience	AIX1	The AI customer service provided product information that was clear and helpful.	Chen et al. (2023); Hsu & Lin (2023)
	AIX2	The AI customer service responded to my questions in a timely manner.	Chen et al. (2023); Hsu & Lin (2023)
	AIX3	The AI customer service effectively helped me resolve questions during the shopping process.	Chen et al. (2023); Hsu & Lin (2023)
	AIX4	The advice provided by the AI customer service matched my actual needs.	Chen et al. (2023); Hsu & Lin (2023)
	AIX5	Overall, I considered this AI customer service experience to be high quality.	Chen et al. (2023); Hsu & Lin (2023)
Perceived diagnosticity	DIA1	After interacting with the AI customer service, I was better able to judge whether this product was suitable for me.	Shin et al. (2013); Lu & Chen (2021)
	DIA2	The AI customer service helped reduce my uncertainty about product choice.	Shin et al. (2013); Lu & Chen (2021)
	DIA3	The AI customer service made it easier for me to compare differences between products or specifications.	Shin et al. (2013); Lu & Chen (2021)
	DIA4	The AI customer service increased my confidence in making a purchase judgement.	Shin et al. (2013); Lu & Chen (2021)
Trust	TRU1	I believe that the information provided by the brand through AI customer service is reliable.	Morgan & Hunt (1994); Huang et al. (2024)
	TRU2	I trust that the brand provides truthful and useful advice through AI customer service.	Morgan & Hunt (1994); Huang et al. (2024)
	TRU3	After interacting with the AI customer service, I had greater trust in the brand.	Morgan & Hunt (1994); Huang et al. (2024)
	TRU4	For purchase-related questions, I am willing to rely on the AI customer service information provided by the brand.	Morgan & Hunt (1994); Huang et al. (2024)
Satisfaction	SAT1	Overall, I was satisfied with this AI customer service interaction.	Oliver (1999); Hsu & Lin (2023)
	SAT2	This AI customer service experience met my expectations.	Oliver (1999); Hsu & Lin (2023)
	SAT3	This AI customer service experience was pleasant.	Oliver (1999); Hsu & Lin (2023)

Construct	Code	Item	Source
Brand loyalty	SAT4	If I experienced it again, I would still consider this service satisfactory.	Oliver (1999); Hsu & Lin (2023)
	LOY1	When purchasing similar cosmetics in the future, I will give priority to this brand.	Oliver (1999); Chen et al. (2023)
	LOY2	I am willing to purchase products again in this brand's live-streaming room.	Oliver (1999); Chen et al. (2023)
	LOY3	I am willing to recommend this brand to others.	Oliver (1999); Chen et al. (2023)
	LOY4	Even if other similar brands are available, I would still tend to continue choosing this brand.	Oliver (1999); Chen et al. (2023)

Note. All items were measured on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree). Items for each construct were adapted, with contextual wording changes, from the source(s) listed in the Source column.