Guardrail-Aware Measurement of Generative AI Marketing in Bangladesh: Consumer Insight, Learning, and Decision Quality

Ruhul Amin
Correspondence: ruhulaminprottoy99@gmail.com

In progress 2026

ABSTRACT

Generative artificial intelligence is accelerating marketing production while creating measurement risks when persuasive efficiency is assessed without consumer trust, recognition, and decision calibration. This study develops a guardrail-aware measurement framework for generative AI marketing in Bangladesh and demonstrates it through a secondary mixed-methods reanalysis of two data streams: 563 organic social-media comments and 131 survey responses. Consumer insight was assessed through sentiment, information-seeking, linguistic variation, and explicit AI callouts; learning was operationalized as recognition and search signals; decision quality was assessed as trust-aligned purchase-intention calibration. Organic comments were predom- inantly neutral, but explicit AI callouts were disproportionately negative. The survey respondents reported high AI detectability, low trust, and a 15.8% purchase-intention reduction under AI-authored advertising. Segment contrasts revealed heterogeneous sensitivity, especially across gender, income, and language groups. The framework extends marketing analytics by treating credibility erosion, verification burden, and social amplification as guardrail indicators alongside conventional response metrics for accountable deployment de- cisions.

Keywords: GenAI Marketing; Guardrail-Aware Measurement; Consumer Trust; Decision Quality; Social Media; Bangladesh

Media; Bangladesh.

1 Introduction

Generative artificial intelligence (GenAI) has evolved from a novelty to a fully functional marketing infrastructure for creating copy, synthesizing images, personalizing content, ideating, engaging with customers, and optimizing campaigns [1]. Recent marketing scholarship presents this shift as a change in production economics and/or value creation for consumers, and the strategic benefits as the ability to scale content and adapt quickly [2]. Previous research on AI in marketing predicted that it would bring automation in segmentation, prediction, communication, and customer management and identified three types of AI – mechanical, analytical, and affec- tive – in research, strategy, and action. Promotional material becomes synthetic by adding GenAI to the mix, and the origin and fidelity of the marketing stimulus becomes part of the consumer evaluation, not a backstage production issue [3].

This shift presents a measurement problem. Traditional digital marketing dashboards focus on reach, im- pressions, engagement, click-through, acquisition, conversion, and return on ad spend. These metrics measure behavioral response, but do not necessarily reflect whether or not the synthetic creative undermined perceived au- thenticity, raised the verification effort, misrepresented the product, or led to poorly calibrated trust. Consumer- AI research highlights the tension between value and discomfort in cases where AI systems collect, categorize, assign or mimic social interactions [4]. When the optimization goals are narrower than the human outcomes that surround them, marketing applications are also plagued by poorly defined goals, lack of transparency, bias, and controllability problems [5]. GenAI exacerbates the tension as the output can be easily manipulated at scale to be persuasive. There is already evidence of the substantial influence potential of personalized persuasive messages and of how marketing systems can be applied to enable automated personalization [6]. A response-only metric can thus be problematic because it may incentivize persuasion in the short run but fail to capture the costs of credibility down the stream.

Research on authenticity and disclosure further focuses the concern. When GenAI is used to replace hu- man creativity, there has been a correlation between a decrease in brand authenticity and a decrease in follower engagement with social media content. The authorship of AI also has an impacts on reactions to marketing communication, especially in terms of authenticity judgements when the messages have emotional and/or rela- tional significance. Different disclosure effects are observed depending on the task that is given to AI, suggesting that consumers do not equate all uses of AI, but rather differentiate between machine placement and machine creation [7]. Trust, sacrifice, convenience, and personalization are also part of that customer-experience research evaluation process [8]. All of these studies suggest that seeing is believing, and that simultaneous measurement of persuasion, recognition, credibility, and information load on the consumer is necessary for effective measurement.

Bangladesh is a significant case in that respect. Social platforms are not just advertising channels, they are also a form of quasi-retail that brings with them all the elements of

advertising: user-generated commen- tary, price questions, product inquiries, and social endorsement. In Dhaka, younger consumers found to be associated with search activity and online purchasing with user-generated and micro-celebrity content. More broadly, Bangladeshi studies also link the promotion features on social media platforms and online reviews to shopping habits [9]. However, most of this literature focuses on social commerce or AI mediated shopping, not on the reactions of the shoppers to the AI-generated promotional material. The national setting also provides linguistic diversity within the Bangla, English, Banglish and code-mixed expressions, which can be identified as sub-audiences that can be easily lost in a national average [10].

Governance is a part of measurement. Engagement totals do not differentiate between failure modes of syn- thetic content: a great-looking photo might actually be a poor representation of a physical feature, a disclosure might actually be about where it came from and not about how convincing it was, or a high interaction post may actually be a result of consumer confusion instead of a successful persuasion. Diagnostic measurement is not simply reporting, but one of these possibilities that makes it a responsible deployment.

The question of whether or not GenAI advertising increases or decreases purchase intent is not an easy one to resolve. Studies that already exist focus on experimental exposure, general AI acceptance, authenticity, or the disclosure effects of AI, whereas discussions about technical guardrails focus on system reliability and not on marketing measurement. In an emerging social-commerce market, little work has been done to combine nat- uralistic consumer commentary with survey-based recognition, trust, and purchase-intention measures. Fewer studies identify a measurement architecture that protects against optimizing persuasion alone through consumer learning and consumer decision quality [8].

The study seeks to answer these questions in four research questions.

• RQ1: What are the consumer insight around social media advertising in Bangladesh based on sentiment, information seeking and AI callout and language signals?

• RQ2: What are the characteristics of consumer learning signals in terms of AI detectability, aesthetic anomaly recognition, trust, and exposure intensity?

• RQ3: How the purchase intention will change as a result of an AI-ad condition overall and by salient customer segments?

• RQ4: What are the appropriate combination of indicators that can be used to provide useful guardrail alerts for marketing measurement?

The goal is to design and validate empirically a Guardrail-Aware Marketing Measurement framework connecting consumer insight, recognition-oriented learning, trust calibration and decision quality. It is methodologically and managerially: it redefines marketing performance as a joint evaluation of response and credibility, and is clear that the data available are aggregate and can be used descriptively to reanalyze, but are not suitable for causal estimation.

2 Literature Review

2.1 Generative AI as a Marketing Production and Persuasion System

GenAI is reshaping the way marketers measure their marketing efforts, as content creation, personalization, and engagement are now on the same computational production floor [11, 12, 13, 14, 15, 16, 17, 18, 19]. The current frameworks talk about flexible inputs, human augmentation, and the speed of output variation, whereas the work that is consumer-focused differentiates between predictive AI and generative systems that produce promotional material directly [17]. Strategic models place AI in research, segmentation, targeting and positioning, and execution [15]. These accounts have a common theme in terms of efficiency, but vary in their focus on consumer side risk [19]. While marketing-pitfall analysis focuses on objective misspecification, opacity, bias and controllability, persuasion research shows scalable psychological tailoring [13]. Applied systems also present practical ways to automated personalization [14]. The main conflict is that a production system that is optimized for response is not necessarily a measurement system that is optimized for informed response [16]. The recognition, authenticity, verification burden and trust diagnostics are therefore necessary for guardrail-aware measurement to be considered a success [18]. There must be a separation between creative productivity, persuasion and consumer facing quality, therefore, measurement is required [12]. If output is increased to the detriment of credibility, there is an unobserved trade-off between output and credibility.

2.2 Authenticity, Disclosure, and Trust Calibration

The most recent studies are most focused on source attribution [20]. When GenAI is used to generate content, it can lead to a loss of authenticity, trust or positive reception if the AI is perceived to be the sole source of the content. However, this impact is different for different tasks. Reactions vary by AI-generated vs. AI-placed ads, and service-ad experiments reveal varying impacts when synthetic generation impacts tangible vs intangible elements. So, it seems consumers are deciding if the AI is involved in representational or relational tasks assigned to it [21].

Research on verification, validation, and trustworthy AI provides a complementary mechanism for understand- ing how consumers can evaluate AI-mediated outputs [22]. Post-hoc explanations have an impact on responses via interpretability and diagnostic value, and explainability is linked to perception, trust, and acceptance in human-AI contexts. Explanation does not ensure calibrated dependency. There must be interaction and friction in interaction design to decrease the likelihood of habitual acceptance, and classes of

explanation vary in their contributions to calibrated trust. Reliability, safety, trustworthiness, and human control are therefore considered as co-goals of human-centered AI. Trust is not enough for marketing measurement, it is important that the reliance is congruent with the consumers' understanding and verification of the marketing stimulus [23] This distinction is between disclosure presence and disclosure quality: authorship is identified by disclosure provenance, and credible claims and usable disclosure verification paths are also a part of calibrated disclosure trust [24].

2.3 Consumer Recognition, Learning, and Decision Quality

Consumer reactions to AI vary by task context. In the algorithm-aversion research, the tasks that are perceived as subjective are found to be less weighted by algorithms, while in the algorithm-appreciation research, the tasks are found to be more weighted by algorithms than human advice in direct comparisons [25]. The contradiction suggests that the machine cue is determined by task fit and perceived human meaning. Customer-experience models also include trust along with convenience, personalization, commitment and sacrifice [26]. Retail rec- ommendation research introduces the concepts of perceived control and choice uncertainty, which are autonomy costs in addition to informational benefits [26].

These streams are also relevant to understanding sustained consumer use of GenAI-enabled shopping sys- tems, where continued engagement depends on more than the immediate purchase outcome [27]. There is no longitudinal recall or comprehension test to test consumer learning and make it a lasting knowledge. It is limited to recognition and information-search signals: perceived detectability, aesthetic anomaly detection and product or price questions. Quality of decision is also not synonymous with increased buying [28]. If the response rate is high after an untrusted creative, it is not necessarily a well-calibrated selection. The quality of the decision is then measured in terms of congruence between intended action and perceived credibility, with the addition of indicators of the effort expended to verify the decision. When combined with information that can be used, recognition thus enables scrutiny, but when combined with synthetic cues, it also reveals the lack of credibility when the cues have not been sufficiently supported.

2.4 Social Feedback, Linguistic Heterogeneity, and Emerging-Market Context

Evaluation, information search, complaint, humor, peer influence, and authenticity monitoring are all aspects of the social media that are not typically observed in controlled exposure studies. Reactions are also visible, and thus form a part of the stimulus for later viewers. Although the focus of GenAI advertising research continues to be authorship, disclosure, personalization, and adoption, emerging-market research also highlights the organiza- tional and contextual conditions that shape AI adoption [29]. This division results in a lack of knowledge about spontaneous skepticism of AI in regular comment sections.

Aggregation is also hindered by the emerging-market context [30]. The language choice, code mixing, income, and platform use are segmentation signals. National averages can obscure sub-audiences that are quicker to pick up on synthetic artifacts or more sensitive to ambiguity. The experience of consumers is also found to depend on the relationship between the person, the task, and the automated system, in consumer-AI research. A Bangladesh- centric approach thus considers the language of the signal as distinct from the language of the commentary, and considers the use of Bangla, English, Banglish and code-mixed commentary as different signal environments rather than different demographic identities. In cases where discovery, evaluation, peer commentary, and seller contact are integrated in the same interface, linguistic segmentation is a tool to identify divergent reactions instead of a demographic term.

2.5 Unresolved Measurement Gap

There are three restrictions along these streams.

1. Studies tend to focus on one of these elements of authenticity, attitude, trust or purchase intention, and a weak endpoint does not indicate which of these elements is responsible for the creation of the authenticity.

2. Experiments provide internal control but provide little unsolicited social control, while observational studies are not typically correlated with structured measures of AI-recognition and purchase intention.

3. Human-centered and explainable-AI research provides guidelines for calibrated trust, but these guidelines are not well translated into campaign analytics.

Thus triangulation over naturalistic comments and survey indicators is necessary, as is explicit segment het- erogeneity, and trust erosion and verification burden as signals relevant to the performance of guardrail-aware measurement. The present secondary reanalysis aims to fill this void without creating models of the causal processes at the level of the individual respondents.

The unanswered question is architectural: How to connect the reaction signals to the governance decisions in a transparent way, without burying the heterogeneous warnings in one score.

3 Framework Conceptualization

The proposed Guardrail-Aware Marketing Measurement (GAMM) framework combines logic for the consumer experience, trust calibration, and human-centered AI. Consumer-AI experience research has found that automated systems influence both instrumental and social/psychological experience [4]. Explainability research reveals that trust is useful when it is used with a level of reliance that is appropriate to the information that the system provides that is understandable to humans [22]. The human-centred AI also

introduces the need for a balance between automation and human control around reliable, safe and trustworthy performance. These concepts are adapted to marketing by considering promotional exposure as a series of measurement states, and not a single persuasion event. The framework is shown in Figure 1.

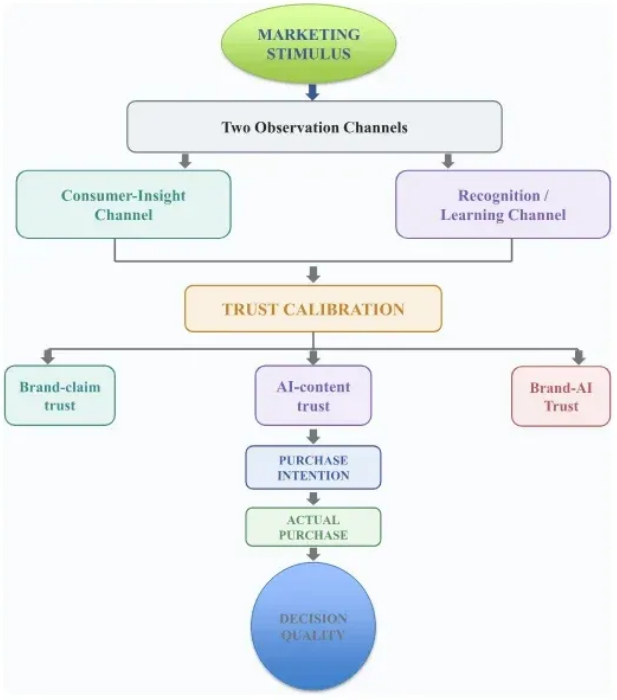


Fig. 1. Guardrail-Aware Marketing Measurement (GAMM) framework

A marketing stimulus is sent to two parallel observation channels, one of which is a GenAI channel. The first is the consumer-insight channel, which includes unsolicited sentiment, inquiry behavior, explicit AI callouts and language-specific reaction. The second is the recognition and learning channel, which includes perceived AI de- tectability, aesthetic anomaly recognition and information search. Both channels contribute to trust calibration, which is baseline brand-claim trust, trust in AI-generated content and trust in brands using AI. Then, the quality of the decision is determined by the extent to which the purchase intention is aligned with the actual purchase, as measured by the difference between the general versus the AI-ad conditions, rather than by the correctness of the purchase itself, as measured ex post.

This pathway is surrounded by four dimensions of guardrails: representational fidelity, provenance trans- parency, consumer agency and social amplification [5]. The framework assumes that a high response rate, but with insufficient credibility, is not a complete marketing response. It also assumes that learning signals are prox- ies of recognition and verification and not a sign of lasting knowledge acquisition. Alerts are visible on the map at a specific location, depending on the segment attributes (language, gender, income, exposure intensity). The analytical output is thus a layered measure of response, with trust erosion, verification burden, heterogeneity and public authenticity policing added to the conventional response. This architecture directly addresses RQ1-RQ4 by placing all of the empirical indicators in a specific mechanism and guardrail function.

The framework does not consider an alert to be evidence of harm, but rather as an opportunity for diagnostic review. The more sources that are combined, the more credible the alert becomes, such as if the anomaly recognition is high, the trust is low, there are negative provenance callouts, and the purchase-intention penalty. The recognition of isolated stability with trust is given less priority. This logic is what makes it possible to not have a strong response metric overwhelm a serious fidelity or trust signal [28].

4 Methodology

Building on the preceding theoretical arguments, Figure 2 presents an integrative conceptual framework that de- lineates the pathways through which GenAI marketing stimuli shape consumer insight, trust calibration, learning, decision quality, and ultimately guardrail-based deployment decisions.

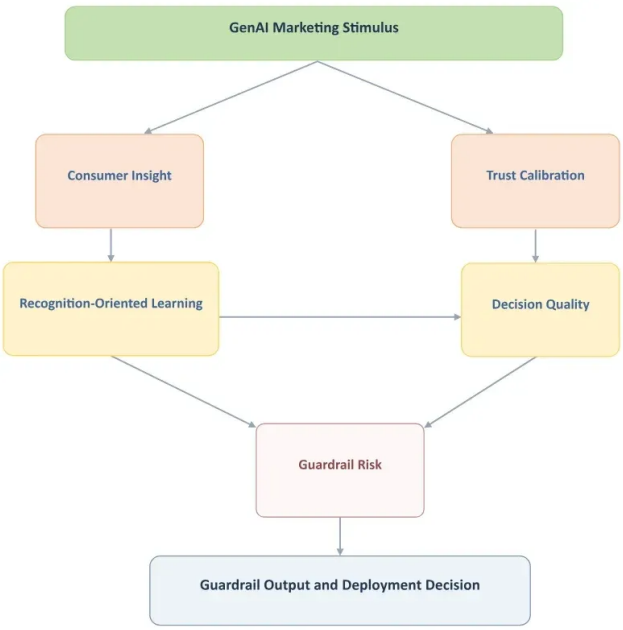


Fig. 2. GenAI Marketing Stimulus–Decision Quality Framework with Guardrail Risk

Taken together, the framework in Figure 2 highlights the interconnected roles of consumer insight, recognition-oriented learning, trust calibration, and decision quality in translating GenAI marketing stimuli into deployment decisions while accounting for potential guardrail risks.

4.1 Research Design and Analytical Logic

The problem of the study was related to the observation of public reaction, as well as the structured evaluation of consumers, for this reason, a convergent secondary mixed-methods measurement design was chosen. The design

was a re-analysis of two real data samples from Bangladesh, one related to comments on social media that were manually coded as organic and the other a structured consumer survey. Streams were analysed separately and the results combined via a common measurement matrix for guardrails. No synthetic respondents were added, no individual records were reconstructed or coefficients were simulated. Building on the preceding theoretical arguments, Figure 2 presents an integrative conceptual framework that delineates the pathways through which GenAI marketing stimuli shape consumer insight, trust calibration, learning, decision quality, and ultimately guardrail-based deployment decisions [31]. So, the analytic unit was coded comment, reported survey construct or published sub-group aggregate.

4.2 Setting, Data Provenance, and Units of Analysis

The context is social media ads and the Bangladeshi consumers, situated within the broader adoption of GenAI-driven digital marketing in emerging markets [30]. The public data were presented in the study ti- tled "AI-Generated Advertising: Consumer Trust and Purchase Intention in Bangladesh" which was presented on Preprints.org manuscript 202608.1207 and date 18 August 2026. A total of 563 unsolicited comments were retrieved from 25 Bangladeshi consumer brands' public Facebook and Instagram pages, which were included in the dataset.The total number of the unsolicited comments in Dataset 1 were 563 (29 from Instagram and 534 from Facebook) from 25 Bangladeshi consumer brands' public Facebook and Instagram pages. There were 131 responses to this survey that rated on a 7-point Likert scale, in Dataset 2. The age range was concentrated in 25–34 years, with male students forming a substantial part of the survey profile; consequently, the results are descriptive of the observed group rather than representative of the wider population, consistent with the importance of socio-demographic heterogeneity in AI-related consumer behaviour [32]. Some branding conclusions were also not made since the majority of comments were contained in a few pages.

4.3 Construct Operationalization

There were 5 domains given prior to reanalysis. Consumer insight included sentiment, thematic inquiry, explicit AI detection and language-specific reaction. This study purposefully focused on recognition-related measures, such as the measures of self-rated AI detectability, aesthetic-appeal deficit, product information seeking, and price information seeking. Trust calibration included baseline trust in social media brand claims, and trust in the AI-generated information versus human-generated information, as well as trust in brands using AI-based advertis- ing. Purchase-intention calibration was the metric used to measure decision quality, not the actual post-purchase utility and accuracy of choice, consistent with the literature's distinction between AI-driven personalization and consumer decision-making outcomes [33]. Purchase-intention calibration was the metric used to measure decision quality, not the actual post-purchase utility and accuracy of choice [34]. The alert indicators for low trust, explicit negative AI callouts, linguistic focus on negativity, penalties for purchase intentions, and exposure related anomaly recognition were used as representations of guardrail risk [35]. The constructs are shown in Table 1 and are related to observable indicators.

As demonstrated in the mapping, recognition, trust and decision calibration are analytically different and no proxy can replace consumer-facing quality.

4.4 Analytical Procedures and Derived Metrics

The aggregated totals and coded percentages of comments for sentiment, thematic structure, AI callouts, and four language categories were used for comment analysis. The means, medians, standard deviations, and subgroup summaries were used for survey analysis. Three derived quantities were calculated which were transparent. Relative purchase-intention change was defined as the difference between the AI-ad mean and mean of the baseline divided by the baseline mean. AI-callout negativity concentration was defined as the ratio of the negative comment's share for AI callouts to the share of negative comments. To see if there was a possible relation between the six reported means of AI detectability and the reported means of relative purchase-intention penalty

Table 1. Guardrail-aware construct operationalization.

Domain Observable indicators Scale / unit Analytical role
Consumer insight Sentiment; thematic inquiry; AI callouts; language segment
Counts and percentages
Detect public uncertainty and social amplification
Recognition- oriented learning
AI detectability; aesthetic deficit; product/price inquiry
7-point scales and percentages
Detect synthetic cues and verification effort Trust calibration Baseline claim trust; AI-content trust; AI-brand trust
7-point scales Compare reliance with perceived provenance and fidelity Decision quality General versus AI-ad purchase intention
7-point means; absolute and relative change
Measure intention calibration under provenance condition
Guardrail risk Negativity concentration; subgroup penalty; segment divergence
Ratios, percentages, ecological association

Trigger review without an opaque composite score by the income-segments, the correlation between these two measures was explored. This Pearson coefficient is an ecological and descriptive one and not an individual level association. Cross-stream triangulation was used to determine if the organic comments and survey indicators were congruent with respect to the dimension of the guardrail.

4.5 Validity, Reliability, and Bias Control

Validity checks were targeted at the overinterpretation. All statistics were rerun from public source data and arithmetically reworked. Original p-values were not reproduced as the original pairwise observations were not available. To maintain the differences between naturalistic comments and survey self-reported, a single scale was not forced. Language was not considered as a demographic identity, a communication segment. Directional interpretation was done with small groups of income. Platform bias and source-concentration bias were identified as the four dominant brand pages and the mix of commentary that was Facebook-centric. Unfortunately the sample was skewed towards younger, male, and student respondents and was therefore not representative. Since the samples of the comments and the surveys were not connected, an agreement across the streams was treated as triangulation instead of an intra-personal agreement.

4.6 Ethics and Reproducibility

Only aggregate, public and de-identified data was used in the reanalysis. The identities of commentators and respondents to surveys and of individual brands from the patterns of their comments were not sought for, nor were verbatim comments from social sites reproduced. The analytical process can be reproduced from the public source providing summary statistics. All values are plotted, as well as displayed in a source-data table next to each plotted value, and all chart objects have embedded worksheets that can be edited. The readability of the arithmetic and the traceability of the aggregates then lie at the heart of reproducibility, not with inaccessible model outputs. Table 2 is an overview of the two data streams, their analytical applications and the main limitations in both data streams. Naturalistic reactions are complementary, not interchangeable: streams of public signals from naturalistic reaction streams, and streams of structured perception and decision measures from survey aggregates.

Table 2. Data streams and analytical coverage.

Data stream N / unit Variables Analytical strength Principal constraint
Organic social-media comments
563 comments; 25 brand pages
Sentiment, theme, language, explicit AI callout
Naturalistic unsolicited reaction
Facebook-heavy and brand-concentrated stream
Consumer survey 131 respondents
Detectability, aesthetic deficit, trust, purchase intention, demographics
Structured common-scale indicators
Nonrepresentative sample; aggregate statistics only
Cross-stream integration
2 independent streams
Guardrail domains and directional convergence
Triangulates public reaction with stated evaluation
No person-level linkage between streams

5 Result Analysis

5.1 Consumer Insight Quality: Sentiment Structure and Inquiry Burden

Consumer insight quality denotes whether public reactions expose evaluation, uncertainty, and information needs hidden by engagement totals; it matters for RQ1 because active commenting is not equivalent to confident acceptance. Table 3 presents the comment-signal matrix. ChatGPT said: Absolutely — here is the same clean, publication-ready format, using booktabs and tabularx so the interpretation column wraps properly.

Table 3: Consumer-insight and verification signal matrix.

Signal Observed value Guardrail interpretation

Evaluative comments 34.8% Direct approval, review, complaint, or support Informational comments 33.7% Active information search around the offer Non-substantive comments 31.4% Background social engagement without clear evaluation Product + price inquiries 29.5% Substantial decision-support and verification burden Explicit AI callouts 2.8% Low-prevalence spontaneous provenance recognition Negative among AI callouts 62.5% High-intensity conditional authenticity concern AI-callout negativity concentration

3.70× overall negative share

Conditional AI skepticism exceeds general negativity

Neutral comments formed the largest share at 55.8%, compared with 27.4% positive and 16.9% negative. Product and price inquiries together represented 29.5%, showing that comment sections also served verification and shopping-support functions. Explicit AI callouts were rare at 2.8%, yet 62.5% were negative. That conditional negativity was about 3.70 times the corpus-wide negative share, making sparse recognition a high-intensity alert. Figure 3 illustrates overall sentiment to separate the dominant neutral information environment from approval and rejection. The result shows why negative-comment frequency alone is incomplete: much of the observable consumer work occurs through neutral verification rather than overt complaint. The thematic mix was nearly balanced across evaluative, informational, and non-substantive content. Net sentiment was +10.5 percentage points, yet the provenance-risk pocket remained concentrated.

5.2 Linguistic Signal Heterogeneity

Linguistic heterogeneity denotes variation across Bangla, English, Banglish, and code-mixed reactions; it matters because pooled averages hide concentrated authenticity risk. Figure 4 compares sentiment composition across the four segments. English comments had the highest positive share at 40.5%. Banglish had the lowest negative share at 2.3% and

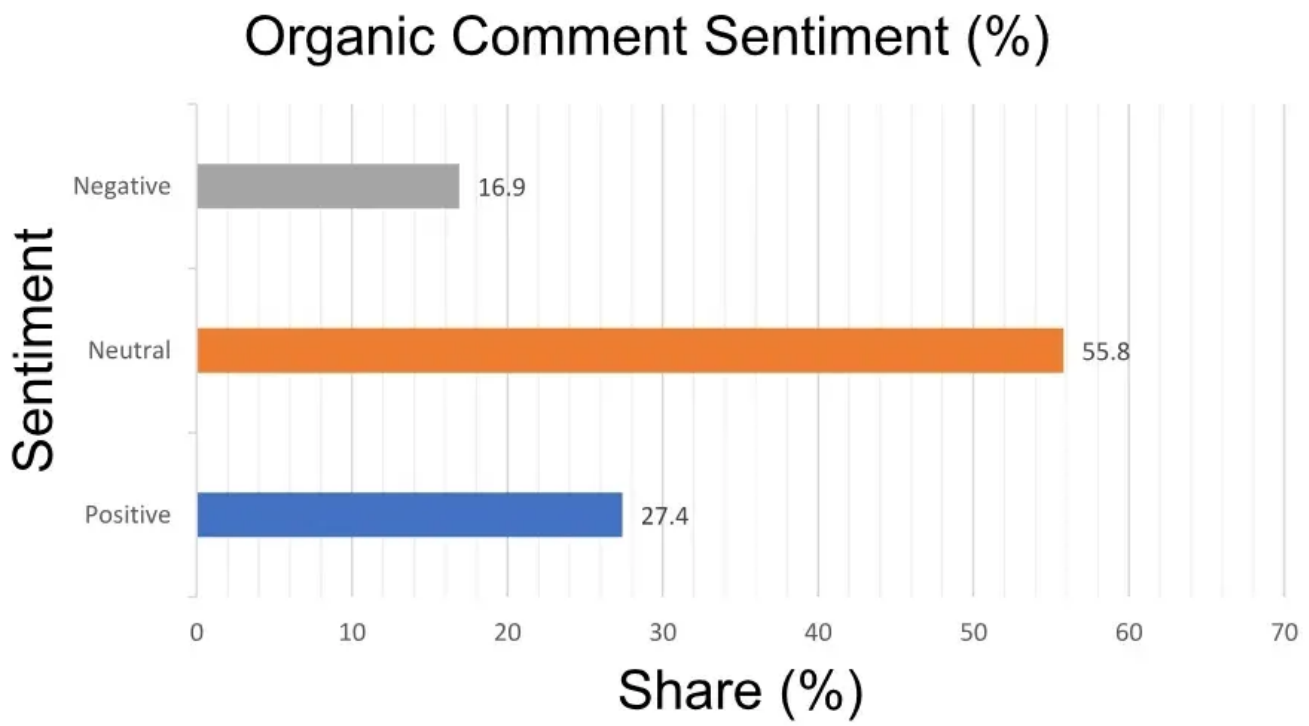


Fig. 3. Overall organic-comment sentiment distribution

the highest neutral share at 79.1%. Code-mixed comments were the outlier: negative sentiment reached 43.5%, equal to the neutral share and roughly 2.57 times the corpus-wide negative rate. Bangla remained closer to the overall profile at 15.0% negative. The pattern identifies concentrated social-feedback risk rather than a uniform national response. Code mixing is treated as a communication segment, not a demographic identity. Segment monitoring therefore surfaces warning signals that an all-comments average suppresses. English net sentiment was +23.7 percentage points versus -30.5 in code-mixed comments, strengthening the case for language-specific monitoring thresholds.

0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%

5.3 Consumer Learning and Recognition Signals

Consumer learning is defined as recognition and information-search behavior rather than durable knowledge; this matters for RQ2 because the data measure perception and inquiry, not retention. Table 4 presents the profile. AI detectability averaged 5.02 on a seven-point scale and aesthetic deficit averaged 4.66. Trust was lower: baseline trust in brand claims was 2.82, trust in AI-generated content 2.80, and trust in AI-using brands 3.08. Heavy social-media users reported an aesthetic-deficit score of 5.08 versus 4.48 among moderate users. Figure 5 illustrates the core survey means on a common scale. Detectability is highest and trust in AI-generated content lowest. This separation indicates recognition without corresponding credibility. Under GAMM, recognition is therefore informative as an alert signal rather than a favorable learning outcome by itself. The detectability-to-AI-trust gap was 2.22 points, while heavy users' aesthetic deficit exceeded moderate users by 0.60, placing recognition upstream of credibility.

Table 4. Recognition, trust, and dispersion indicators.

Indicator Median / reported level SD / contrast Measurement interpretation
AI detectability Median 5.00 SD 1.73 Recognition is high but dispersed Aesthetic-appeal deficit Median 5.00 SD 1.96 Visual anomaly perception is heterogeneous Trust in AI-generated content
Median 3.00 SD 1.70 Trust remains below the scale midpoint Trust in AI-using brands
Median 3.00 SD 1.57 Brand-level trust remains constrained General purchase intention
Median 4.00 SD 1.58 Baseline intention centers on the midpoint AI-ad purchase intention
Median 4.00 SD 1.73 Dispersion rises under synthetic provenance Screen-time contrast Moderate aesthetic deficit 4.48

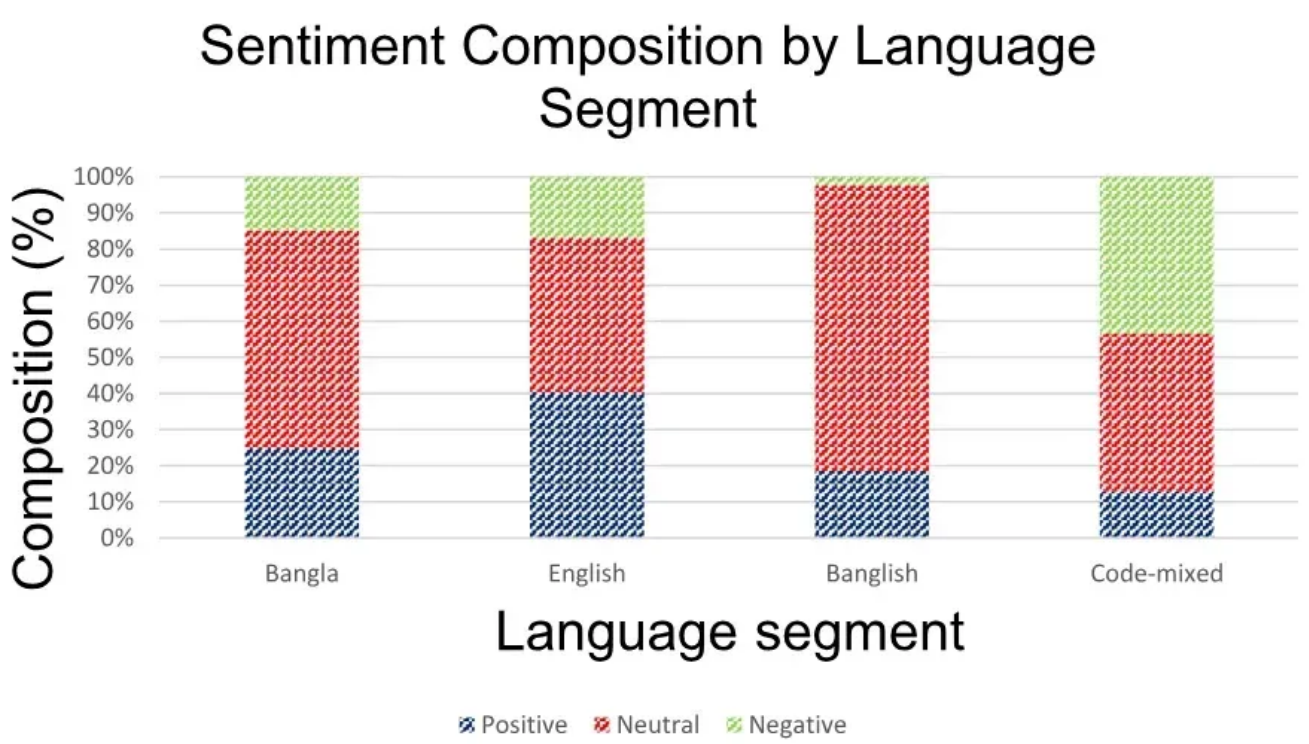


Fig. 4. Sentiment composition across language and dialect segments

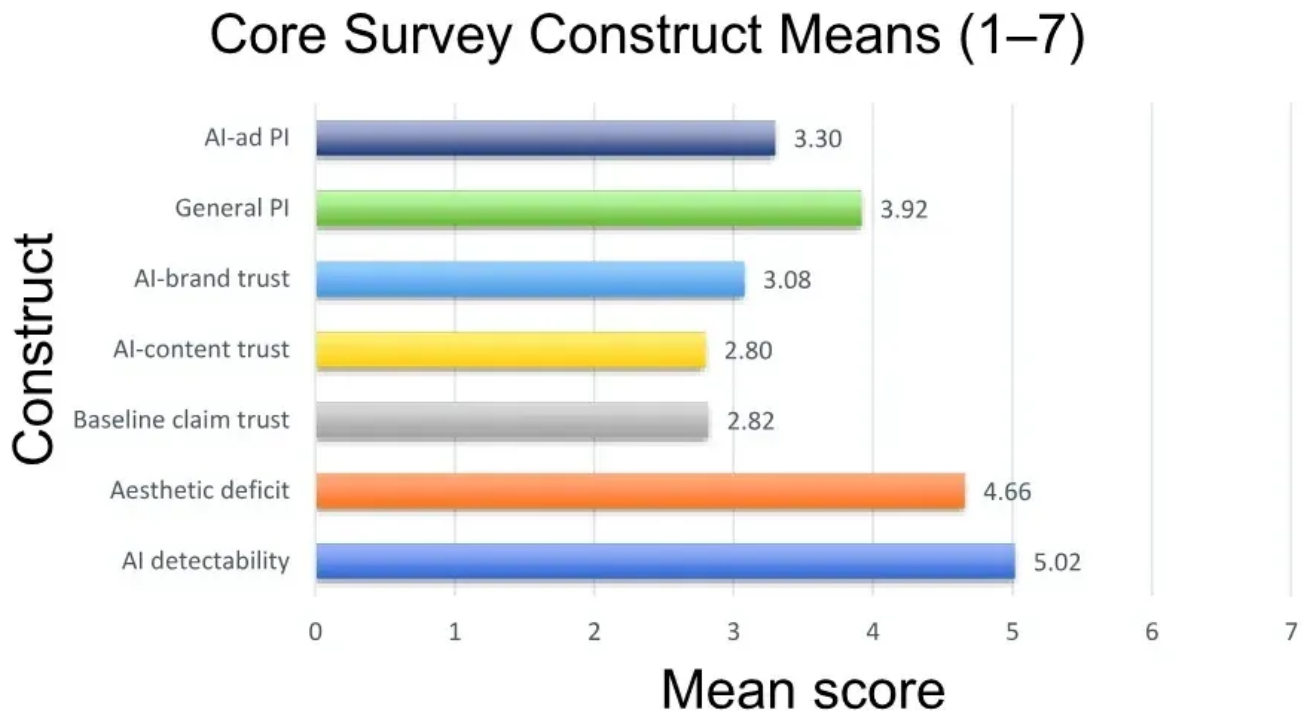


Fig. 5. Core survey construct means on seven-point scales

5.4 Decision Quality under the AI-Ad Condition

Decision quality denotes calibration of purchase intention to advertising credibility; it matters for RQ3 because synthetic provenance changes the decision state even when the marketed product is unchanged. Table 5 presents overall and gender contrasts.

Table 5. Purchase-intention calibration by gender.

Group	N	Absolute change (AI–baseline)	Relative change	Retention of baseline
Female	33	−1.28	−27.9%	72.1%
Male	98	−0.40	−10.8%	89.2%
Combined	131	−0.62	−15.8%	84.2%

Combined purchase intention declined from 3.92 to 3.30, a 0.62-point or 15.8% relative penalty. Female respondents moved from 4.58 to 3.30, a 27.9% decline, whereas male respondents moved from 3.70 to 3.30, a 10.8% decline. Figure 6 compares baseline and AI-ad purchase intention for both groups and the combined sample.

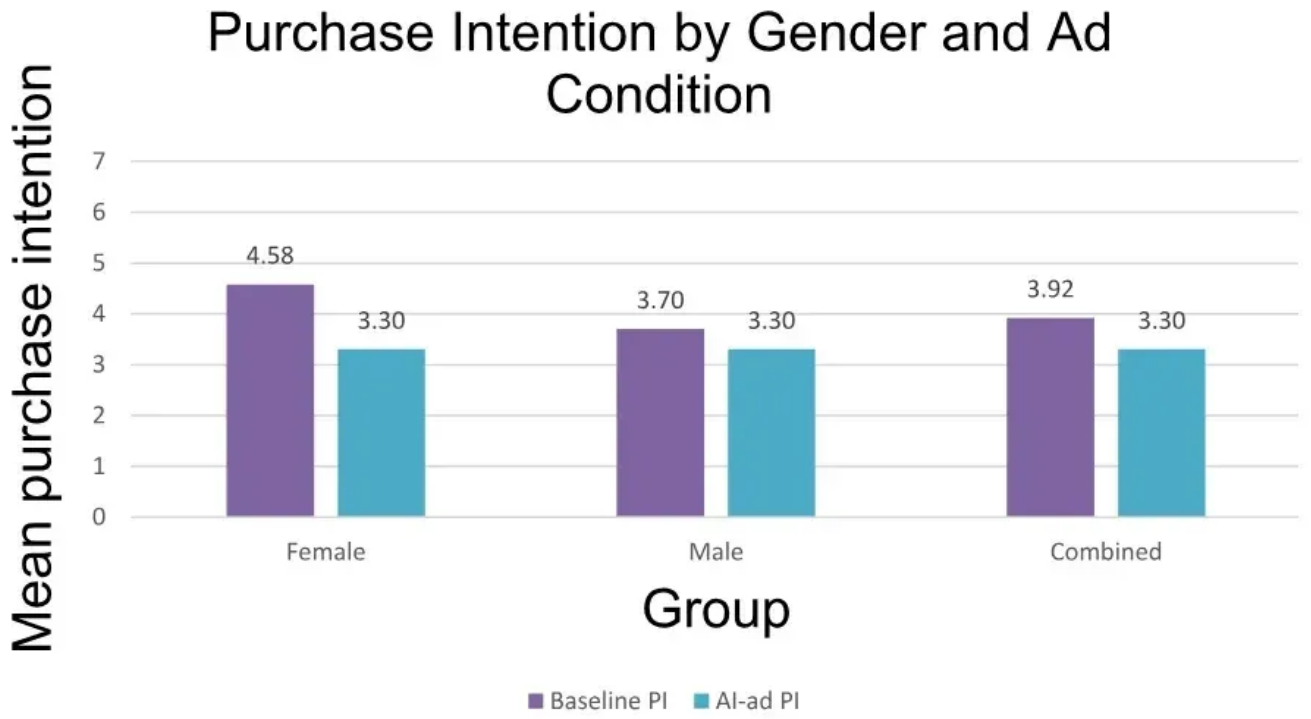


Fig. 6. Purchase intention by gender and advertising condition

The female subgroup shows the strongest loss, while AI-ad means converge at 3.30. The result is not inter- preted as a universal gender mechanism; it demonstrates that an overall mean hides materially different sensitivity within this sample. Guardrail reporting therefore retains subgroup deltas beside the aggregate response met- ric. Purchase-intention retention was 84.2% overall, 72.1% for women, and 89.2% for men; the female baseline advantage disappeared under the shared AI-ad mean.

5.5 Segment Sensitivity and Ecological Guardrail Risk

Segment sensitivity asks whether groups with stronger AI recognition also show larger purchase-intention penal- ties; it matters for directing monitoring without individual-level inference. Figure 7 plots mean detectability against relative purchase-intention penalty for six income groups.

4.3 4.55 4.8 5.05 5.3 5.55 5.8Relative PI penalty (%) Mean AI detectability (1–7)

The 50,000-100,000 BDT group combines the highest detectability, 5.84, with the largest relative penalty, 25.93%. Student or no-income respondents show the smallest penalty, 9.38%, at detectability 4.65. The six-point ecological correlation is weakly positive at r = 0.276. The scatter rejects a simple monotonic rule. Similar de- tectability levels coexist with different penalties, while the 10,000-24,999 BDT group shows a substantial penalty despite the lowest detectability. Recognition is therefore one guardrail dimension, not a sufficient explanation of decision change. Income penalties ranged from 12.20% above 100,000 BDT to 22.59% at 10,000-24,999 BDT outside the reported extremes, confirming no simple gradient.

5.6 Integrated Guardrail-Aware Measurement

Integrated guardrail measurement combines heterogeneous warnings into a decision-facing profile; it addresses RQ4 by identifying signals that complement engagement or conversion metrics. Table 6 presents the scorecard across fidelity, provenance, trust, verification burden, segmentation, and decision calibration.

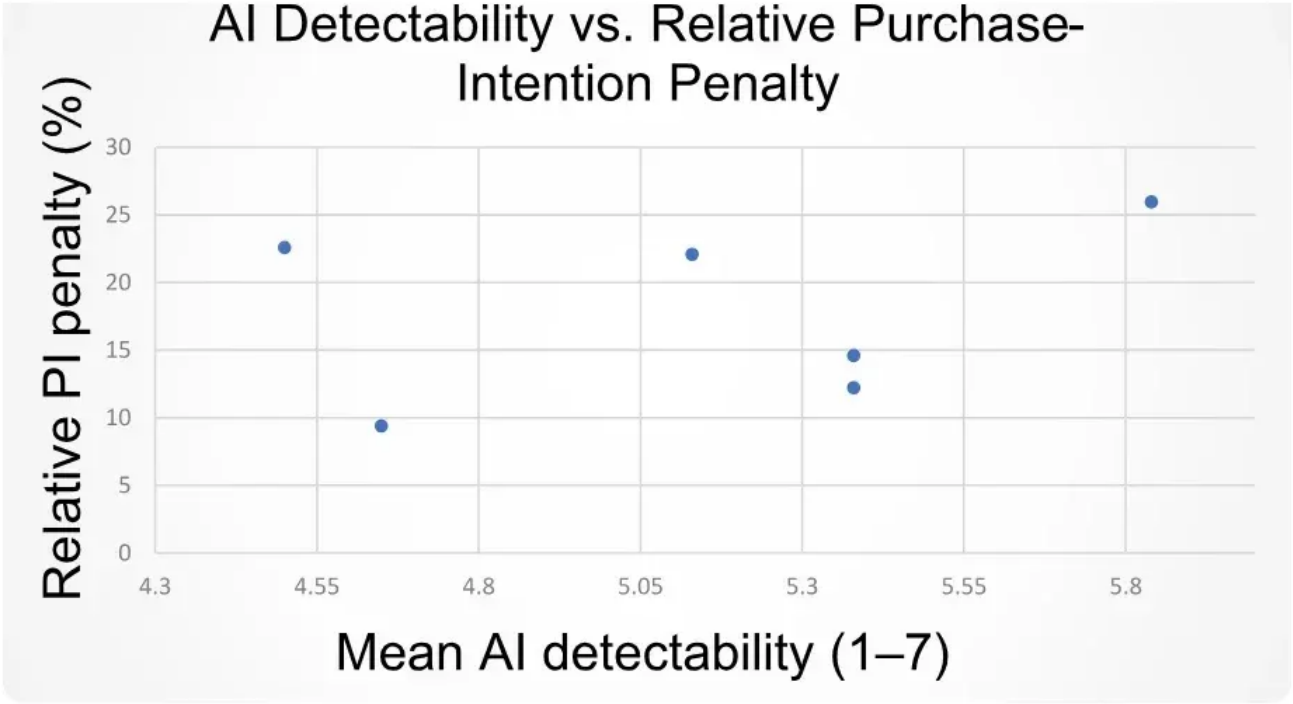


Fig. 7. Income-segment AI detectability versus relative purchase-intention penalty

Table 6. Integrated guardrail-aware measurement scorecard.

Dimension Observed alert Status Deployment implication
Representational fidelity
Aesthetic anomaly perception is materially elevated
Elevated Human review for truth-bearing product visuals Provenance transparency
AI callouts are sparse but conditionally negative
High-intensity, low-frequency
Track provenance callouts and thread escalation Trust calibration Trust remains below the scale midpoint
Elevated Pair provenance disclosure with verifiable details Verification burden Information seeking occupies a material share of comments
Material Strengthen specifications, price clarity, and product evidence Segment heterogeneity
Language, gender, and income signals diverge
Elevated Use segmented monitoring thresholds with minimum sample rules Decision calibration AI-ad condition reduces stated purchase propensity
Elevated Compare provenance-condition deltas before wider deployment

No composite index is calculated because the indicators differ in scale and meaning. Figure 8 illustrates five percentage-based alert rates: explicit AI callouts, negativity among those callouts, product-or-price inquiry burden, code-mixed negative sentiment, and overall purchase-intention penalty.

Negativity among AI callouts is highest at 62.5%, followed by code-mixed negativity at 43.5%; explicit callout prevalence is lowest at 2.8%. The contrast shows why frequency alone is inadequate. Rare signals become operationally important when conditional negativity is extreme, while inquiry burden signals broader verification work. GAMM therefore retains response metrics but escalates diagnostic review when credibility, verification, or segment alerts deteriorate. The scorecard therefore uses escalation rather than compensation: severe fidelity or trust alerts remain visible despite strong engagement, especially when conditional negativity is concentrated. These layers convert scattered reactions into reproducible governance signals.

6 Discussion

The results back up this measurement interpretation of GenAI marketing, where the outcomes of persuasion and credibility are linked, but not identical. Neutrality and information seeking were the predominant sentiments in the organic comment stream, and there were no explicit calls for AI but a few negative ones. Interestingly, the sur- vey also revealed that they were highly self-reportedly detectable, had visible aesthetic deficits, did not trust the AI-ad condition, and had even lower purchase intentions under the condition. The patterns are consistent with the GAMM framework, in that recognition and public feedback precede trust calibration, and the quality of deci- sion is measured by whether intended action is in line with credible information and not just because of exposure.

The authenticity mechanism is particularly consistent with experimental evidence that brand authenticity is lower when GenAI is used to create social media compared to when it is created by humans and that there are

Figure 8: Percentage-based guardrail alert rates; indicators retain distinct meanings and do not form a composite index fewer positive responses when the brand creates marketing

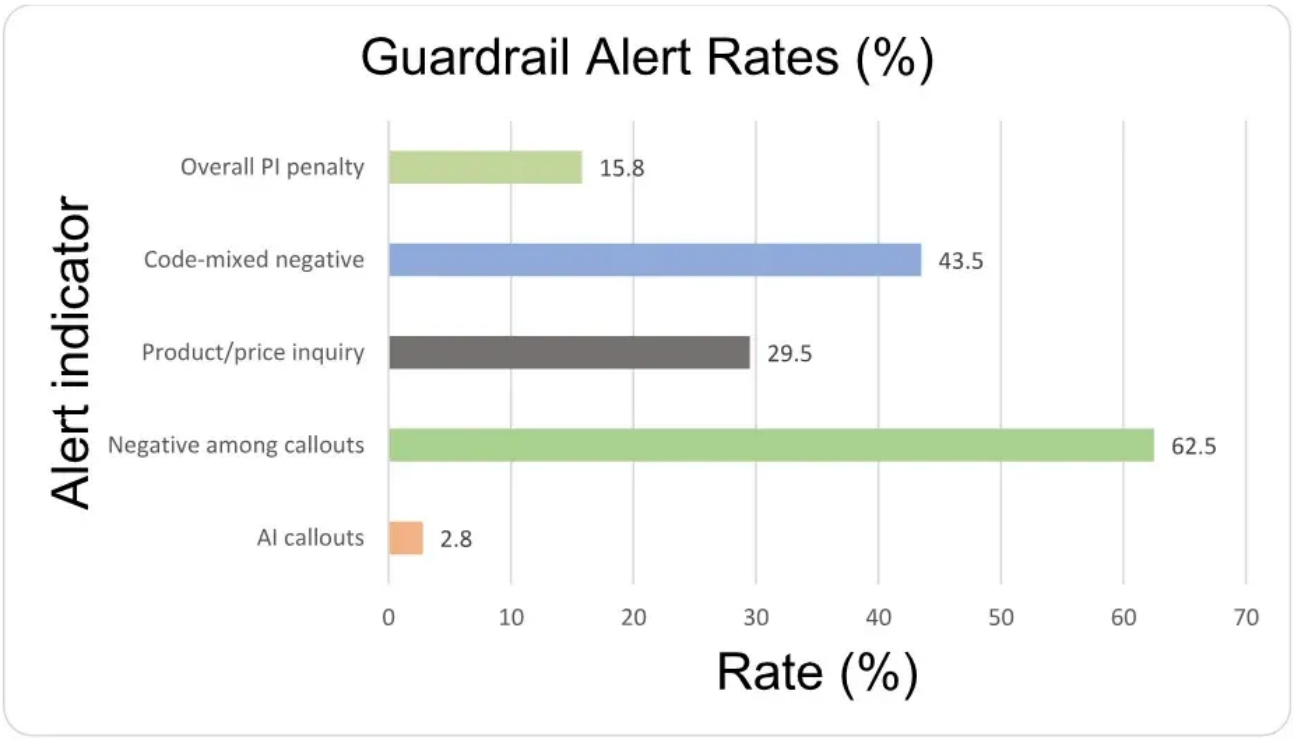


Fig. 8. Percentage-based guardrail alert rates; indicators retain distinct meanings and do not form a composite

communication through AI vs humans' authorship. The Bangladesh reanalysis takes that line of thought a step further, and provides a naturalistic social-feedback environment. While few of them explicitly referenced AI, the mean score of those that did was more negative than the average. That implies that authenticity policing is an event which is socially amplified: the authenticity challenges which are made visible by a handful of comments can shift the interpretive setting for later viewers – even if the majority of comments are not AI related.

Disclosure research: fills in some needed details. AI disclosure is not always harmful: its impact on trust is mediated by its influence on the tangible and intangible components of the service representation suggests that there are different types of AI disclosure effects and that the effects of AI disclosure on trust are mediated by the effect of AI disclosure on the tangible or intangible components of the service representation. The current results corroborate that conditional perspective as detectability does not translate uniformly to loss in purchase intentions among income groups. It is therefore more appropriate to think of recognition as an alert trigger, rather than as a cause of the issue. WEC and heterogeneous subgroup penalties suggest baseline trust and expectations, as well as perceived fidelity and task appropriateness are potential intervening processes.

The low trust profile also supports the notion of the difference between transparency and calibrated re- liance. Explainable recommendation research reveals that post-hoc explanations affect consumer responses via interpretability but that there is a difference between increasing the amount of explanation and increasing the amount of calibration. From a marketing point of view, this implies that attributes like provenance labelling, product specification, unperturbed views of the product, transparency in pricing, and the ability to check the authenticity along with the label must be measured. An honest statement with not-so-imaginary artwork doesn't solve the fidelity issue.

The results of the decisions are of managerial importance. The overall purchase-intention penalty of 15.8% is significant because it is in the context of already low levels of trust in normal claims made by social media brand advertisers. The above-mentioned non-monotonic income pattern and the bigger female size reduction also demonstrate the importance of national averages for the governance of campaigns. Recommendation-agent research also includes a relationship of technology-assisted decisions with a decrease in perceived control and choice uncertainty. Click and conversion is not the only attribute that a guardrail-aware dashboard can benefit from—deltas in segments, inquiry rates, trust measures, and authenticity call-outs can all help.

The contribution is methodical and deliberate, its restraint. Secondary data can be used to make clear com- parisons and triangulations are possible, while mediation and causal attributions at the individual level are not possible. The framework thus omits the reconstruction of structural coefficients and considers the ecological association descriptive only. This restriction restricts inference, and enhances verifiability. The key takeaway is that measuring GenAI marketing in Bangladesh as a system of consumer insight, recognition-oriented learning, trust calibration, and

decision quality provides more meaningful insights for marketing—than measuring each end-point of a marketing campaign individually.

There are also policy implications to the framework. There are other layers to reveal, those of consumer facing guardrails, requiring fidelity checks regarding attributes that are true of the system, easily accessible verification methods, and monitoring for social amplification following deployment. This is different from the idea of a change of information basis of choice—the content of stylistic synthesis. Research designs should also include results for recognition, calibration of trust, effort expended on the verification, perceived control, and actual behavior. The multilingual monitoring is of special significance in Bangladesh, because there are large differences between code-mixed reactions and the pooled one. Responsible AI is not just about model governance, it also applies to metrics that are employed to assess marketing performance.

7 Conclusion

This study tackled the issue that traditional GenAI marketing measurement only accounts for observable re- sponse and fails to capture credibility, recognition, and verification burden, and decision calibration. A GAMM framework was created and secondary mixed methods reanalysis of 563 organic social-media comments and ag- gregate survey responses from 131 consumers in Bangladesh was conducted. The naturalistic stream had a very neutral environment with a significant amount of product and price enquiry, with very few explicit calls to AI that were negative. The perceived detectability of AI was high, trust was low, and the purchase intention was reduced by 15.8% in the AI-ad condition in the survey aggregates. There were also differences in risk by gender, income, exposure, and language.

The contribution is to view consumer insight, recognition-oriented learning, trust calibration and purchase-intention quality as a chain of measurement points instead of as separate points in the chain. This allows marketing teams to get meaningful alerts on guardrails without turning all the signals into a black box score. The study is still constrained by the use of secondary data that are aggregated, the composition of the surveys is not representative, the concentration of platforms and brands, the lack of linked data streams, and the lack of behavioral conversion measures and longitudinal learning measures. Future studies should include preregistered experiments, representative probability sampling, multivariate analysis at the respondent level, actual conversion outcomes, and repeated exposure designs with a separation of recognition and durable learning and objective decision accuracy.

Recommendations Bangladeshi brands using GenAI advertising need a dual dashboard that pairs campaign-response metrics with credibility guardrails. Creative review needs escalation when synthetic output changes product color, proportion, texture, human likeness, or other truth-bearing attributes. Provenance disclosure needs to be paired with verifi- able product information rather than treated as a standalone remedy. Comment monitoring needs separate flags for explicit AI callouts, product and price inquiries, and language-segment negativity so that low-frequency but high-intensity authenticity concerns are visible early. Survey tracking needs repeated measures of detectability, aesthetic anomaly, trust, and purchase-intention deltas rather than a single satisfaction score. Segment-level re- porting needs minimum sample thresholds and uncertainty notes before operational targeting decisions. Human review needs to remain accountable for high-risk creative assets and claims. For research teams, de-identified microdata, coding protocols, and preregistered analysis plans need to accompany studies so that mediation, subgroup effects, and decision-quality mechanisms are tested directly rather than inferred from aggregates.

Data Availability Aggregate statistics were sourced from Preprints.org manuscript 202608.1207, posted 18 August 2026: https://www.preprints.org/manuscript/202608.1207. No individual records were reconstructed.

References

[1] Dhruv Grewal, Cinthia Satornino, Thomas Davenport, and Abhijit Guha. How generative ai is shaping the future of marketing. Journal of the Academy of Marketing Science, 53:702–722, 2024.

[2] Erik Hermann and S Puntoni. Artificial intelligence and consumer behavior: From predictive to generative ai. Journal of Business Research, pages –, 2024.

[3] Gun Ho Lee, Kyoung Jun Lee, Baek Jeong, and Taekyung Kim. Developing personalized marketing service using generative ai. IEEE Access, 12:22394–22402, 2024.

[4] Simoni Rohden and L Espartel. Consumer reactions to technology in retail: choice uncertainty and reduced perceived control in decisions assisted by recommendation agents. Electronic Commerce Research, 24:901– 923, 2024.

[5] Erik Hermann and Stefano Puntoni. Generative ai in marketing and principles for ethical design and deployment. Journal of Public Policy and Marketing, 44:332–349, 7 2025.

[6] S Matz, J Teeny, S Vaid, H Peters, G Harari, and M Cerf. The potential of generative ai for personalized persuasion at scale. Scientific Reports, 14:–, 2024.

[7] K Lei and Yixuan Liu. When ai becomes a shopping advisor: A study on the impact of generative ai review on consumer purchase decision. SAGE Open, 15:–, 2025.

[8] J Hou, Shuai Yang, Guiyang Xiong, and P Pavlou. Enhancing ai-assisted purchase decisions: The role of the sense of autonomy. MIS Q., 50:589–614, 2025.

[9] Sanjida Nourin, Md. Ashiqur Rahman, Md. Elias Hossain, and M Iqbal. Consumers' attitude toward ai- driven e-commerce adoption in bangladesh: An extension of planned behavior. Bangladesh Journal of Multidisciplinary Scientific Research, pages –, 2025.

[10] Nabila Islam and H Misbah. Digital faith meets sustainability: The role of ai and religiosity in shaping green purchase intentions in bangladesh. In The Proceedings of the ASEAN School of Business Network International Conference, pages –, 2025.

[11] Bsmah Alsulami, Budoor Alwated, Khadijah Barashid, Manal Abdullah, Monerah AlOsaimi, and Samah Alhusayni. On leveraging generative artificial intelligence (genai) for behavior learning and personalized marketing optimization. Archives for Technical Sciences, pages –, 2025.

[12] Syed Hamza Aziz and Muhammad Zohaib Razzaq. Artificial intelligence in digital marketing: Enhanc- ing consumer engagement and supporting sustainable behavior. Journal of Computational Informatics & Business, pages –, 2025.

[13] Omotoso Oluwayomi Joseph, A Akintola, E Offiong, Olaonipekun Olaitan Olajuwon, Udodirim Ibem Offia, and Adedayo Ayodeji Faniyan. Ai-enhanced consumer insights: Leveraging behavioural analytics for hyper- personalised marketing strategies. Journal of Artificial Intelligence, Machine Learning and Data Science, pages –, 2025.

[14] Komal Komal. The role of ai in predicting consumer behaviour and optimizing marketing campaigns. IN- TERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, pages –, 2025.

[15] Prashant Kumar. The role of artificial intelligence in shaping consumer behaviour: A study on ai-driven marketing strategies. International Journal of Business and Management Invention, pages –, 2025.

[16] Mitra Madanchian. Generative ai for consumer behavior prediction: Techniques and applications. Sustain- ability, pages –, 2024.

[17] Mahanish Panda, Kandarp Singh, A Raj, and Roma Puri. Understanding the interplay between generative artificial intelligence and consumer behaviour through a systematic literature review. Discover Artificial Intelligence, 6:–, 2025.

[18] Marko Sarstedt, Susanne J. Adler, and Monika Imschloss. Machine marketing: rethinking the customer in the age of generative ai. Journal of Marketing Analytics 2026, pages 1–11, 7 2026.

[19] Hao Xie, Yuankun Nie, Li Liu, and Yang Chen. Machine learning-based research of ai marketing: topic analysis and model construction. Future Business Journal, 11:–, 2025.

[20] Yi Dong, Ronghui Mu, Yanghao Zhang, Siqi Sun, Tianle Zhang, Changshun Wu, Gao Jin, Yi Qi, Jinwei Hu, Jie Meng, S Bensalem, and Xiaowei Huang. Safeguarding large language models: a survey. Artificial Intelligence Review, 58:–, 2024.

[21] Borga Babadoˇgan. Unveiling the power of ai-driven personalization: Transforming consumer behavior in the age of digital marketing. Next Frontier For Life Sciences and AI, pages –, 2024.

[22] Xiaowei Huang, Wenjie Ruan, Wei Huang, Gao Jin, Yizhen Dong, Changshun Wu, S Bensalem, Ronghui Mu, Yi Qi, Xingyu Zhao, Kaiwen Cai, Yanghao Zhang, Sihao Wu, Peipei Xu, Dengyu Wu, Andr´e Freitas, and Mustafa Mustafa. A survey of safety and trustworthiness of large language models through the lens of verification and validation. Artificial Intelligence Review, 57:–, 2023.

[23] Ming Hui Huang and Roland T. Rust. The genai future of consumer research. Journal of Consumer Research, 52:4–17, 6 2025.

[24] Md Mehedi Hasan Emon. Digital transformation in emerging markets: Adoption dynamics of ai image generation in marketing practices. Telematics and Informatics Reports, pages –, 2025.

[25] Yuejiao Fan and Xianggang Liu. Exploring the role of ai algorithmic agents: The impact of algorithmic decision autonomy on consumer purchase decisions. Frontiers in Psychology, 13:–, 2022.

[26] Yan Guo, Qichao Tang, Haoran Wang, Mengjing Jia, and Wei Wang. Artificial intelligent housekeeper based on consumer purchase decision: a case study of online e-commerce. Ind. Manag. Data Syst., 124:2588–2612, 2024.

[27] Behzad Foroughi, Morteza Ghobakhloo, Jun Wen, and Masood Fathi. Sustained use of generative ai for shopping: a pls-ann analysis. The Service Industries Journal, 46:250–283, 2025.

[28] A Mari, A Mandelli, and Ren´e Algesheimer. Ai-assisted shopping on voice assistants: The effi- ciency–autonomy consumer decision trade-off. Journal of Interactive Marketing, 61:327–344, 2025.

[29] Mohammed Enshassi, R Nathan, Soekmawati, and H Ismail. Unveiling barriers and drivers of ai adoption for digital marketing in malaysian smes. Journal of Open Innovation: Technology, Market, and Complexity, pages –, 2025.

[30] Khawaja Mohammad Mustaqeem. The impact of generative ai on digital marketing strategies: Evidence from emerging markets. International Journal of Research and Scientific Innovation, pages –, 2025.

[31] E Mogaji and Varsha Jain. How generative ai is (will) change consumer behaviour: Postulating the potential impact and implications for research, practice, and policy. Journal of Consumer Behaviour, pages –, 2024.

[32] Rahul Raj Singh Dr. Rama and Arvind Kumar Pandey. Ai and digital marketing adoption in bihar: A study of consumer behaviour across socio-demographic segments. Economic Sciences, pages –, 2025.

[33] Hari Leela.Vemula, Hyderabad Telangana, Daryao Khatri, Academic Coordinator, D Vijayalakshmi, Sachin Hatole, and A India. Artificial intelligence in consumer decision-making: A review of ai-driven personaliza- tion and its managerial implications. Journal of Informatics Education and Research, pages –, 2025.

[34] Wenxin Wang, Zhiguang Chen, and Jiwei Kuang. Artificial intelligence-driven recommendations and func- tional food purchases: Understanding consumer decision-making. Foods, 14:–, 2025.

[35] Zachary Ravichandran, Alexander Robey, Vijay Kumar, George Pappas, and Hamed Hassani. Safety guardrails for llm-enabled robots. IEEE Robotics and Automation Letters, 11:4649–4656, 2025.