Does Structured Grounded Generation Improve Consumer Response to AI-Generated Social Media Posts? Evidence from a Randomized Online Experiment and Implications for Outcome-Based Pricing

Shengan (Lucian) Lu (s3799964)

Supervisors:
Natalia Amat Lefort & Bas Kruiswijk

BACHELOR THESIS– COMPUTER SCIENCE & DATA SCIENCE AND ARTIFICIAL INTELLIGENCE

Leiden Institute of Advanced Computer Science (LIACS)
liacs.leidenuniv.nl

June 27, 2026

Abstract

While small businesses are using generative AI to create marketing materials, there are questions of whether the process of generation, specifically the pipeline rather than the model itself, influences consumer response. The present thesis presents a randomized, between-subjects, preregistered online experiment ($N = 261$ after exclusions) testing two post-generation pipelines that rely on the same base large language model in the area of restaurant social-media marketing. A Single-Shot Generation (SSG) pipeline is compared to a Structured Grounded Generation (SGG) pipeline that enforces a structured template, restricts all factual claims to a grounded scenario card, adds authority and social-proof cues, and checks hard constraints before a post is finalized. Each participant viewed exactly one of the 40 realized stimuli and was asked about their visit/purchase intention (confirmatory primary outcome), click intention, save intention, perceived credibility, informativeness and social-proof strength, as well as authenticity. There was no detectable effect of SGG on visit/purchase intention using the proportional-odds cumulative link model with fixed effects for scenarios and cluster-robust standard errors at the stimulus level (OR = 1.01, 95% CI: $[0.66, 1.55]$, $p = .97$). At the same time, the manipulation was modestly successful—SGG posts were rated as more specific and fact-grounded (manipulation index $d = 0.27$)—and, in exploratory secondary analyses (exploratory because the confirmatory primary test was null), SGG was associated with higher perceived credibility (+0.25 points on a five-point scale, $p_{\text{Holm}} = .003$). These perceptual gains were largely decoupled from behavioral intention: visit intention correlated only weakly with perceived credibility ($r = .01$), informativeness ($r = .08$), and social-proof strength ($r = .05$). An illustrative outcome-based pricing exercise returns a zero variable component because the estimated high-intent uplift (0.003, 95% CI $[-0.13, 0.14]$) is not statistically distinguishable from zero. The results indicate that, absent disclosure during exposure, structured grounded generation can help to give AI-generated posts a sense of credibility without actually shifting the behavioral intent that would be monetized by outcome-based pricing.

Keywords: AI-generated content; large language models; structured grounded generation; consumer response; perceived credibility; randomized online experiment; outcome-based pricing; social-media marketing.

Contents

1 Introduction

Small and medium-sized enterprises (SMEs) are increasingly leveraging generative artificial intelligence (AI) and large language models (LLMs) to generate marketing content and other customer-facing messages [36]. LLM will provide a regular stream of social media posts with near-zero marginal cost for a small restaurant or shop. However, when companies are buying AI-assisted marketing services, they don't care so much about stylistic fluency, as whether the service brings them measurable improvement in consumer response, or "attributable value." This means that the pricing mechanism of such services is key because if, for instance, a price is only based on the volume of output, this can incentivize fluent but ineffective writing, whereas a pricing mechanism based on value or on the outcome will require measurable and verifiable output.

This thesis explores that issue in relation to the concrete use case of restaurant social media posts created using AI. The important methodological step is the distinction between the *generation pipeline* and the *underlying model*. The study does not compare two LLMs, but rather two post-generation processes using the same base model. Single-Shot Generation (SSG) is a process that generates a post using a single prompt. A Structured Grounded Generation (SGG) approach enforces a rhetorical structure, constrains each factual statement to a specified "scenario card", includes cues of authority (establishing facts that are easily verifiable), and cues of social proof (combined scores of reviews), and checks hard constraints before the post is completed. Thus the treatment is the workflow, with a constant model identity. Both types of posts can in principle be made from the same base model, so any differences in consumer response will be due to the engineering around the model and not the model itself.

1.1 Problem statement and research question

The key question is: Does this type of content engineering matter for consumers if the AI-authorship goes undisclosed? The effect of the pipeline is isolated by keeping disclosure fixed [4, 24]. The main research question is:

> *In a randomized online experiment, does a Structured Grounded Generation (SGG) strategy produce more favorable consumer response than a Single-Shot Generation (SSG) strategy?*

The study addresses four sub-questions:

1. (Confirmatory) Does SGG increase visit/purchase intention relative to SSG?
2. (Secondary behavioral) Does SGG increase click intention and save intention?
3. (Secondary perceptual) Does SGG increase perceived credibility, informativeness, and social-proof strength, and how does it relate to perceived authenticity?
4. (Translation) How can an experimentally estimated response uplift be translated into an illustrative pricing exercise for AI-assisted marketing services?

1.2 Contributions

This thesis makes three contributions. First, it offers a mechanically clean (as opposed to a model effect) estimate of a *pipeline* effect on consumer response, based on a stimulus-sampling design with 40 actual posts, which reduces the risk of a confounding of the treatment effect with the idiosyncratic characteristics of one or two posts in particular [47]. Second, it records a dissociation: the degree to which a post is perceived as credible and specific and fact-based is increased by the SGG pipeline without the headline behavior (visit/purchase) being altered. Third, it is done via an explicit illustrative outcome based pricing exercise that illustrates how, in principle, this evidence can be mapped into a transparent pricing rule and why, in this dataset, this rule leads to a zero price for the variable component.

In short, the confirmatory behavioral outcome—visit/purchase intention—was not affected by the SGG pipeline: there was no detectable difference (OR = 1.01, $p = .97$). It did consistently shift how posts were perceived, however—raising perceived credibility in exploratory analyses (+0.25 points, $p_{\text{Holm}} = .003$)—but not behavior, with the result that the illustrative outcome-based pricing rule charges nothing for its variable component. The contribution is thus not a positive content-quality effect, but rather a robust, pre-specified null together with this dissociation and its pricing implication.

1.3 Thesis overview

This chapter presented the problem and research questions. Section 2 presents the theoretical framework (Elaboration Likelihood Model as the primary lens, signaling theory and the MAIN model as the secondary lenses, and the Persuasion Knowledge Model as the boundary condition), the estimands and hypotheses. In Section 3 related works on AI generated marketing material and consumer response is discussed. The methodology, or experimental design, stimuli, measures, and analysis plan is described in Section 4. The results for the realized sample are provided in section 5. In Section 6 the results are discussed based on the theoretical framework and the threats to validity are addressed. The contribution and future research is summarised in Section 7. This is a bachelor thesis that was done at the Leiden Institute of Advanced Computer Science (LIACS) in the supervision of Natalia Amat Lefort and Bas Kruiswijk.

2 Background

Here, the stimuli are not consumer-generated word-of-mouth, but marketer-generated persuasive posts, so although this research on electronic word-of-mouth (eWOM) is informative with regards to credibility and purchase intention, it does not reflect the word-of-mouth of consumers. Thus, eWOM is considered adjacent literature that supports the credibility and social-proof cues [21], while social-proof is viewed as a signal-like cue in brand communication, the persuasive effect of which may be impacted by the concreteness, credibility, and groundedness of its appearance, not by its inherent authenticity as a peer-to-peer testimonial [19, 41]. The thesis focuses on three complementary perspectives to explain why the SGG pipeline may impact consumers' response.

2.1 Primary framework: the Elaboration Likelihood Model

More specific, informative, and well-organized arguments should facilitate the central-route processing of a message, thereby enhancing the quality and relevance of the message arguments, whereas formatting, fluency, and source cues offer peripheral cues [7, 38]. The SGG pipeline aims at improving both routes: The grounded factual structure and claim-to-evidence mapping target argument quality (central route) while improved formatting, structural markers and embedded credibility cues target peripheral processing. There are two restrictions to ELM that apply. First, the central–peripheral dichotomy can be confounded in fast-paced social-media environments where much information is processed in the periphery, as noted above [25]. Second, ELM is descriptive rather than predictive, it identifies processing conditions, but does not specify quantitative effect sizes. These restrictions constrain the power of theory-based predictions, but not the framework.

2.2 Secondary framework I: signaling theory

Signaling theory, developed in Spence's analysis of the job market [41], provides a rationale for why the specific design features of SGG should improve consumer response. Under information asymmetry in markets, credible signals help to mitigate uncertainty and enhance exchange [9, 41]. As an example, restaurant marketing is a canonical example where the consumer cannot determine the quality of food, the quality of service and the veracity of promotional claims without visiting the restaurant. Signaling theory has been used in digital commerce to design websites, to use specificity of product information, and also to add third party endorsements and verifiable claims as quality signals to decrease perceived risk [32, 48]. The central idea in the prediction is that signals must be hard to mimic: a claim which is specific and verifiable is more effective than promotional language that is more general [9]. This is similar to the contrast between SSG and SGG, which validates claims, incorporates authority and social-proof cues, and enforces factual grounding, all of which boost signal specificity and verifiability. Signaling theory is important because the consumer need not identify the signaler (or the signal) which renders it appropriate to a no-disclosure design.

2.3 Secondary framework II: the MAIN model and technology heuristics

The thesis is also informed by Sundar's MAIN model of technology mediated credibility [42] which considers four categories of heuristics that drive the assessment of digital content: Modality, Agency, Interactivity, and Navigability. Two heuristics under the Agency dimension are especially pertinent: *Authority cues* (establishment year, neighborhood history etc.) are facts that invoke authority-related credibility judgements [29, 42]; *Social-proof cues* (aggregated review scores etc.) represent the behavior of others and invoke bandwagon-like inferences [29, 42]. The two are working on different pathways of inference, source expertise and peer consensus and SGG works on both. The MAIN heuristics, unlike the Persuasion Knowledge Model below, can be used without knowledge of the origin of the content, making the framework suitable for an unlabeled-post design [43].

2.4 Boundary condition: the Persuasion Knowledge Model

The Persuasion Knowledge Model (PKM) states that consumers can cope with persuasion attempts by knowing the persuasion agents' goal and tactics [14]. Recent extensions have theorized that consumers might use different coping strategies when they see the persuasion agent as AI [45] but that can only happen if consumers perceive AI as a source – exactly what this no-disclosure design holds constant. This means that, in the future disclosure-rich environments, in particular under the transparency requirements of the EU AI Act and the forthcoming European Commission transparency guidance [5, 12, 13, 28],[1] PKM is viewed as a boundary condition and not as a primary mechanism.

2.5 Treatment definition and causal logic

Since the study is a comparison between two pipelines constructed from the same base LLM, the treatment is not model identity but instead the workflow. The SSG is a single-shot generation workflow. SGG implements a content structure, only allows factual claims in a specific scenario card, and checks hard constraints prior to the finalization of the post, in line with the general NLP philosophy that explicit grounding leads to better traceability in fact and to less content without supporting evidence [26]. The treatment effect is interpreted as the causal effect of the SGG pipeline (as a whole), not of each individual component (structure, grounding, validation, authority cues, or social-proof cues) alone. With regard to theoretical perspectives, all three mechanisms (signaling, ELM, MAIN) suggest that SGG should generate more favorable consumer response compared to SSG.

2.6 Estimands and hypotheses

Given that the primary outcome (visit/purchase intention) is an ordinal variable, the primary confirmatory estimand is the proportional-odds treatment parameter β for SGG in the cumulative-logit model defined in Section 4.6 (or equivalently, the cumulative odds ratio $\exp(\beta)$ by which exposure to SGG as opposed to SSG increases the odds of being in a higher versus a lower response category). Confirmatory hypotheses are directional ($H_1 : \beta > 0$ or OR > 1), and 2-sided 95% confidence intervals and p-values are reported. The descriptive and linear robustness check of the mean difference $\mathbb{E}[\mathit{VisitIntent} \mid \mathit{SGG}] - \mathbb{E}[\mathit{VisitIntent} \mid \mathit{SSG}]$ is reported, but is not the confirmatory estimand. The secondary hypotheses are related to perceived credibility, informativeness, and social proof strength. Authenticity is a secondary (and exploratory) outcome due to mixed literature that is highly disclosure sensitive [1, 4, 24].

3 Related Work

According to recent marketing research, AI authorship and AI disclosure can lead to negative responses by decreasing perceived authenticity, trust, or advertising value. GenAI's role in the content creation aspects of social-media branding can erode the authenticity of a brand, although there are also indications that the negative impact is less pronounced when AI is used as a supplement to human content creation processes [4]. Beyond this, when consumers consider content for emotional marketing as AI-generated, there can be negative word-of-mouth and a drop in consumer loyalty [24]. The same issues are reported regarding restaurant branding [1], service advertising [16] and disclosed AI-generated advertising [5, 28]. However, it is a challenging test case, not an easy one, as the concern is increased in hedonic context, where consumers are less receptive to AI recommenders compared to utilitarian contexts [30].

[1] The cited Commission document is the second draft published on 5 March 2026, which is the guidance on transparency. Later final guidance may differ from the guidance in this draft, as the Article 50 transparency guidance is continuing to develop at the time of writing.

But there are many negative impacts of AI-authorship based on disclosure or source awareness. If consumers are not able to reliably tell the difference between AI-generated and human-generated text, the source-based reactions might not be the same [22]. The present study aims to investigate whether the *pipeline-quality* differences matter, when there is no disclosure, but to discuss the effects of disclosure by AI as a boundary condition. It also relates to a methodological literature on stimulus sampling in experimental persuasion research, which finds that when only one or two stimuli are used per condition, the observed effect could be due to idiosyncratic aspects of the specific stimuli rather than the manipulation [23, 47, 49]. The design below is a solution that realizes five stimuli per scenario-by-condition cell and explicitly models stimulus identity.

4 Methodology

The study is a randomized, between-subjects online experiment using Qualtrics with online recruitment of participants via student groups. Each subject was exposed to only one of the experimental conditions and only one stimulus post. The one-post-per-participant design eliminates carryover and direct comparison effects that might occur if the same person observed SSG and SGG.

4.1 Design structure

Four restaurant scenarios (to ensure non-trivial stimulus variety within one domain) were crossed with two conditions (SSG and SGG) and five retained valid stimuli per scenario-by-condition cell for the experiment. This provides 40 realized post stimuli ($4 \times 2 \times 5$). Each participant viewed one post only. Realizing five stimuli per cell, instead of one or two, directly tackles the stimulus-sampling problem [47] — in that case, the random intercept for stimulus identity in the analysis model has a meaningful distribution to estimate, and the treatment effect is less susceptible to confounding with stimulus-specific features.

4.2 Definition of stimuli and treatment

Each scenario comprised a fixed image, a structured scenario card with all the factual information available to the model (type of cuisine, item highlighted, price, location, opening hours, promotion rule, authority fields, social-proof fields and prohibited claims), and the post variants generated under both SSG and SGG. Two distinct types of credibility-relevant information were identified on the scenario card—those that are verifiable facts about the business (e.g., the founding year) and those that are aggregated consumer ratings (e.g., "4.4 stars from 5,200+ Google reviews").

The base LLM used for both conditions was the same, therefore the treatment was the generation pipeline rather than the model itself. In the **Single-Shot Generation (SSG)** condition, the posts were generated using the regular single shot prompt, but importantly, the authority and social-proof fields were not provided to SSG, which only received basic information about a restaurant. The **Structured Grounded Generation (SGG)** condition involved generating the posts based on the scenario card, but using a structured, grounded, constraint-validated workflow that enforces a pre-specified template (hook, short context, value points, evidence-based social-proof cue, call-to-action), does not allow for claims beyond the material from the scenario card, and includes a validation step to check for unsupported claims, prohibited phrases, and internal inconsistencies.

For a failed draft, a single repair round was allowed. Candidate outputs were generated sequentially until five valid stimuli were collected for each cell; no manual stylistic editing was done to the first five valid outputs. To enhance the ecological validity, the stimuli were stylized as realistic mobile screenshots of social media posts, with the same non-treatment visual elements (e.g. profile name, profile image, time stamp, engagement icons) kept constant across each scenario. The two pipelines are summarized in Figure 1 and a realized SSG/SGG stimulus pair for one of the scenarios is shown in Figure 2.

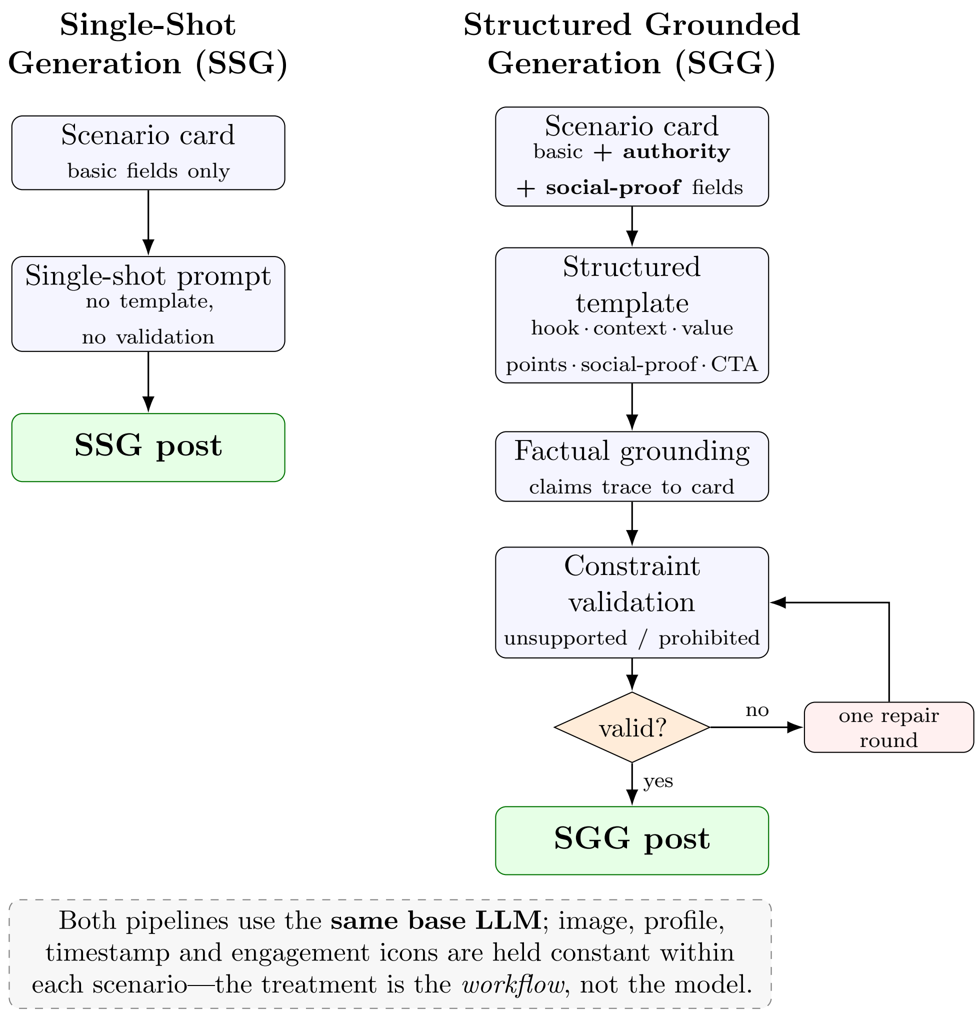

Figure 1: The two post-generation pipelines. SSG produces a post from a single prompt using only basic restaurant fields. SGG uses the same base model but additionally receives authority and social-proof fields and applies a fixed rhetorical template, factual grounding to the scenario card, and a constraint-validation step with at most one repair round. The estimated treatment is the bundled SGG workflow relative to SSG.

(a) SSG

(b) SGG

Figure 2: A realized SSG/SGG stimulus pair for the same scenario, with the image, profile, and engagement elements held constant. Relative to SSG, the SGG caption adds an *authority cue* ("…serving New York-style slices …since 1975 …"), a *social-proof cue* ("Rated 4.4 stars by 5,200+ Google reviewers"), explicitly structured value points, and a clear call-to-action—the credibility-relevant features whose pooled effect the experiment estimates.

4.3 Participants and procedure

The participants were adults (18+) and a convenience sample was recruited online via student group messaging and the survey was conducted in English. The convenience sample is mentioned as a limitation in generalizability, however, care was taken with data quality in this mode of recruitment, through the attention-check and speeding exclusions described below. The participants were randomized to the eight cells of the scenario-by-condition matrix and then to one of the five stimuli retained within a cell after providing informed consent. They filled in the dependent variables immediately after exposure, followed by the manipulation-check items (inserted after the main outcomes), one attention-check item, and background questions. The participants were debriefed at the end and told that the posts were artificially produced by AI. The anticipated analytic sample ranged from 260-300 usable responses while recruitment resulted in 261 usable responses (see Section 5.1 for exclusions). The statistical power is therefore sufficient for moderate effects, but not as high for small effects and this is considered a caveat.

4.4 Measures

All evaluative responses were measured on 5-point scales, and where possible, were worded to match those used in previous advertising and social media research [3, 17, 31, 39]. The confirmatory primary outcome was visit/purchase intention (1 item); secondary behavioral outcomes were click intention and save intention (both 1 item), secondary perceptual outcomes were perceived credibility (3 items, adapted from Ohanian [37]; credibility is a multi-faceted construct, see Diamantopoulos et al. [11] for more detail), perceived informativeness (1 item; acceptable for concrete attribute, see Bergkvist and Rossiter [2] for more detail), perceived social-proof strength (1 item), and perceived authenticity (3 items, adapted from Napoli et al. [35]); manipulation checks were whether the post seemed well organised, specific/concrete, and fact grounded (all 1 item); background covariates were Instagram-use frequency (1 item), dining-out frequency (1 item), and general attitudes towards AI.

4.5 Exclusion rules

After preregistration, answers were only discarded if the primary outcome was not completed, duplicated, failed the attention check or were unusually fast. Under the speeding rule, any response with a completion time below the larger of 90 seconds or one-third of the median valid completion time was excluded. Due to the fact that there was no pilot completion time for this analysis, the median from the main sample was used as a proxy, thus setting the threshold at 90 seconds.

4.6 Analysis plan

The main confirmatory estimand is the intent-to-treat effect on visit/purchase intention of SGG versus SSG. The outcome is ordinal and so the preregistered primary model is a cumulative link mixed model (CLMM) with treatment as a fixed effect, a set of fixed effects for the different scenarios, and a random intercept for realized stimulus identity:

$$\operatorname{logit}\{\Pr(Y_i \leq k)\} = \tau_k - \beta \cdot \mathbb{1}[SGG]_i - \alpha_{\text{scenario}(i)} - u_{\text{stimulus}(i)}.$$

The pre-registered first choice was this CLMM (e.g. R's `ordinal` package [8]) and the pre-registration specified a fallback for the case of a singular or otherwise unavailable mixed fit: a cumulative link model (CLM) with scenario fixed effects and cluster-robust standard errors at the stimulus level [6, 34]. As it turned out, a mixed-ordinal version was not available in the analysis environment and this pre-registered fallback was instead employed: proportional-odds cumulative logit with treatment and scenario fixed effects, estimated by maximum likelihood in Python's `statsmodels` software package with a cluster-robust (sandwich) covariance at the stimulus level. A second random-intercept logistic model on the dichotomised high-intent outcome was used as a diagnostic for the omitted random intercept, and the between-stimulus variance was found to be small (Section 5.3) warranting the use of the cluster-robust CLM. Primary outcome was assessed at the .05 level without multiplicity adjustment; secondary outcomes were assessed with the Holm procedure within the pre-specified secondary family, gated on the primary outcome test [18]. A linear model using HC3 robust standard errors is provided as a robustness analysis as suggested by [15, 27]. Scenario-specific estimates, exploratory behavioral index, covariate moderation, mediation-consistent associations are presented as exploratory, with the intention–behavior gap as standing caveat for all intention outcomes [40, 46].

4.7 Ethics

The participants in the study were adults and the survey itself was low risk and conducted online. Participants were asked to opt-in and consent to participate in the program. As AI writing was not revealed on exposure, delayed disclosure was used which was addressed by end-of-study debriefing. No platform account identifiers were taken and there is no personal identification of participants in the data.

5 Results

The results on the realized sample are reported in this chapter. Confirmatory analyses (primary outcome) are distinguished throughout from secondary and exploratory analyses.

5.1 Sample, exclusions, and descriptives

From the 321 responses collected, 60 responses were excluded under the preregistered exclusion rules, resulting in an analytic sample of $N = 261$ (Table 1). The dominant exclusion was incomplete responses (missing the primary outcome, $n = 56$); three responses were a failure of the attention check and one a speeder (4 seconds - the next-fastest valid response took 117 seconds). The analytic sample of $N = 261$ agrees with the proposed sample size range of approximately 260 to 300 useable responses, which supports sensitivity to moderate effects but leaves limited sensitivity to small effects, as described in Section 6. Seven of 261 retained responses were marked as incomplete by the survey site, but they all had complete data for every analysis variable, and the same substantive pattern is obtained in a sensitivity analysis restricted to the 254 fully completed responses (visit-intention OR = 0.97, 95% CI $[0.63, 1.49]$, $p = .89$).

Table 1: Sample flow under the preregistered exclusion rules.

Stage	n
Collected responses	321
Excluded: incomplete (missing primary outcome)	−56
Excluded: duplicate	−0
Excluded: failed attention check	−3
Excluded: speeding ($< 90\,\mathrm{s}$)	−1
Analytic sample	**261**

The analytic sample included 123 in SSG and 138 in SGG, from a total of 40 stimuli (median 7 responses per stimulus, range 2–8) and four scenarios ($n = 63$–70 per scenario; cell sizes 29–38). Respondents were generally frequent social-media users (mean Instagram-use frequency 2.46 on a 1=*multiple times a day* to 5=*do not use* scale) and eating out frequently (mean = 2.52). Overall attitudes towards AI were slightly positive to neutral (mean = 2.77 on a scale from 1 = very positive to 5 = very negative). Cronbach's α was 0.84 for the credibility scale, 0.81 for the authenticity scale, and 0.83 for the manipulation-check index, indicating that the three multi-item scales had good internal consistency [10].

Means, standard deviations, and standardized differences (Cohen's d) for each outcome in each condition are presented in Table 2. The pattern previews the main result: the behavioural outcomes (visit and click) change very little, while the perceptual outcomes (credibility, informativeness and social-proof strength) improve by about a fourth of a standard deviation for SGG. The distributions of the primary outcome in each condition are virtually identical as seen in Figure 3.

Table 2: Descriptive statistics by condition ($N = 261$; all measures on 1–5 scales). Positive d favors SGG.

	SSG		SGG			
Outcome	M	SD	M	SD	Diff.	d
Visit/purchase intention (primary)	3.47	0.81	3.49	0.78	+0.01	0.02
Click intention	3.56	0.58	3.54	0.56	−0.03	−0.04
Save intention	3.33	0.70	3.46	0.75	+0.14	0.19
Perceived credibility	3.45	0.57	3.70	0.70	+0.24	0.38
Perceived informativeness	3.38	0.67	3.55	0.71	+0.17	0.24
Perceived social-proof strength	3.33	0.67	3.51	0.75	+0.19	0.27
Perceived authenticity	3.31	0.61	3.35	0.55	+0.04	0.08

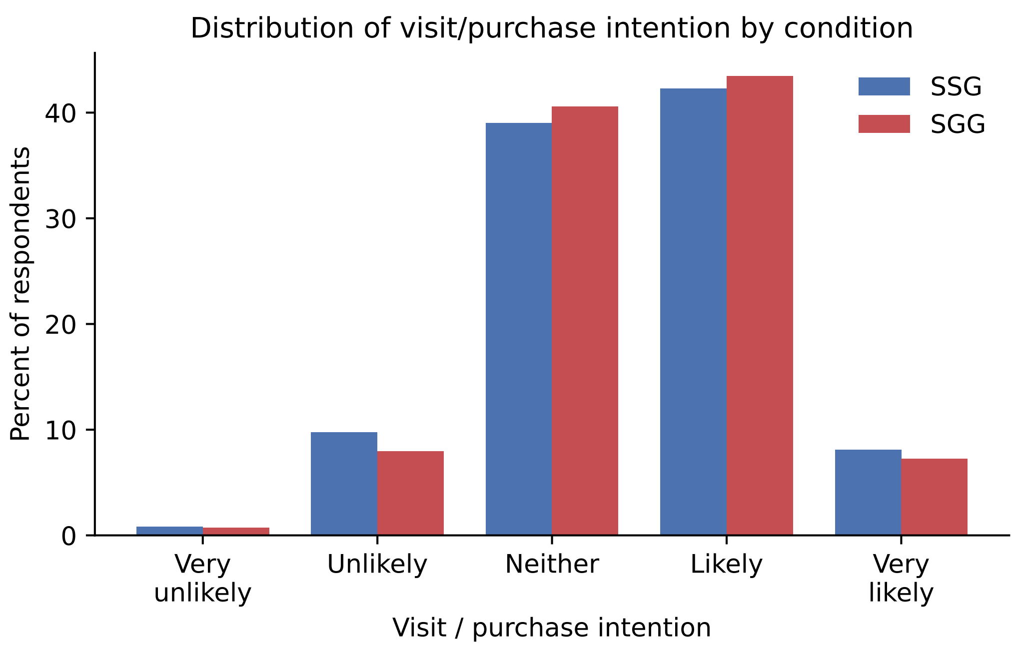

Figure 3: Distribution of the primary outcome (visit/purchase intention) by condition. The two distributions are nearly indistinguishable.

5.2 Manipulation checks

Manipulation proved to be modestly successful. Relative to SSG, SGG posts were rated as more specific and concrete (3.82 vs. 3.63, $d = 0.26$) and more fact-grounded (3.82 vs. 3.65, $d = 0.26$), with a smaller difference on being well organized (3.76 vs. 3.63, $d = 0.19$). The combined manipulation index was higher for SGG (3.80 vs. 3.64, $d = 0.27$, Mann–Whitney $p = .035$). The SGG pipeline has indeed altered the intended perceptions of structure and grounding, but to a small extent (Table 3, Fig. 4). This has significance for interpretation: a null effect on the primary outcome cannot be interpreted as having a completely failed manipulation.

Table 3: Manipulation-check ratings by condition.

Item	SSG M	SGG M	Diff.	d	Mann–Whitney p
Well organized	3.63	3.76	+0.14	0.19	.171
Specific / concrete	3.63	3.82	+0.19	0.26	.035
Fact-grounded	3.65	3.82	+0.17	0.26	.055
Manipulation index (mean)	3.64	3.80	+0.16	0.27	.035

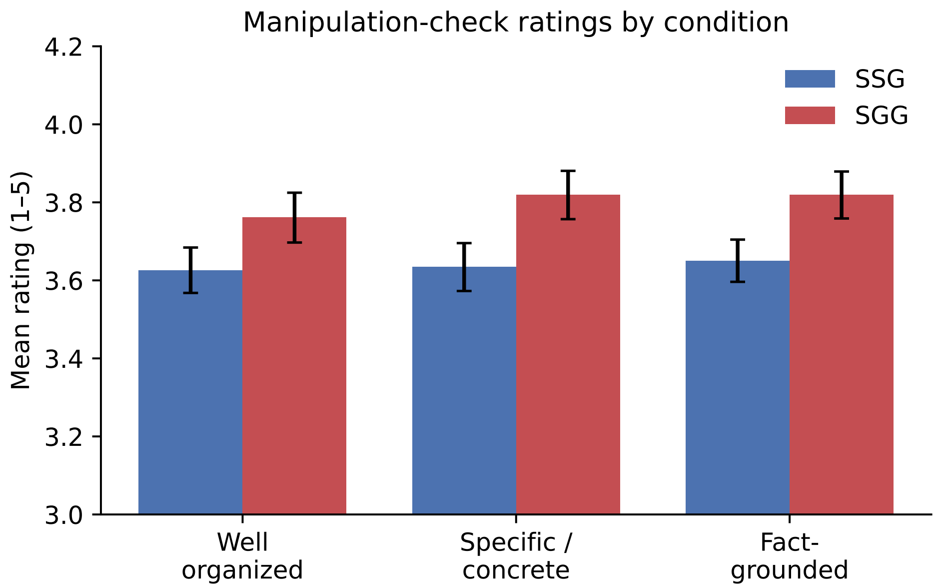

Figure 4: Manipulation-check ratings by condition (means ±1 SEM). SGG posts were perceived as somewhat more specific and fact-grounded.

5.3 Primary confirmatory analysis

No significant differences were found between SGG and SSG in visit/purchase intention. The stimulus-level cluster-robust SGG cumulative odds ratio was 1.01 (95% CI $[0.66, 1.55]$, $p = .97$) for the proportional-odds cumulative logit (with scenario fixed effects). The point estimate is very close to no change in the latent intention distribution and the confidence interval excludes any effects that are more than a moderate OR either way. A cluster-robust standard error estimated as 0.22 on the log-odds scale was close to the model-based standard error estimated as 0.24 and a supplementary variational-Bayes logistic model on the dichotomized high-intent outcome estimated a small stimulus-level random-intercept standard deviation (≈ 0.23). The cluster-robust CLM is pre-registered and was used for this analysis because a mixed-ordinal implementation was unavailable in the analysis environment, although such diagnostics suggest that the variance between stimuli is small, meaning that the absence of a random intercept is unlikely to materially affect the conclusions.

Since the confirmatory primary test was not significant, the pre-registered testing hierarchy suggests the secondary family should be investigated in an exploratory fashion, and the secondary Holm adjusted tests below are presented for completeness but should not be interpreted as confirmatory.

5.4 Secondary outcomes (exploratory)

Table 4 reports the secondary family. Within the exploratory secondary family, perceived credibility remained Holm-adjusted significant ($p_{\mathrm{Holm}} = .003$) and was associated with a 0.25-point higher rating on the five-point scale under SGG. Perceived informativeness (OR = 1.49) and social-proof strength (OR = 1.74) were nominally higher under SGG at the raw level ($p = .047$ and $p = .033$) but did not survive multiplicity adjustment. The behavioral secondaries—click and save intention—and perceived authenticity showed no reliable difference. Figure 5 summarizes the estimated SGG−SSG differences across all outcomes on the common 1–5 scale.

Table 4: Secondary outcomes (exploratory; the confirmatory gate was not passed). Effect is the cumulative odds ratio for ordinal outcomes and the cluster-robust mean difference for the credibility/authenticity composites. p_{Holm} within the six-member secondary family.

Outcome	Effect type	Estimate	95% CI	p_{raw}	p_{Holm}	Sig.
Perceived credibility	mean diff.	0.25	$[0.11, 0.39]$	.0006	.003	yes
Perceived social-proof strength	odds ratio	1.74	$[1.05, 2.88]$	.033	.165	no
Perceived informativeness	odds ratio	1.49	$[1.01, 2.19]$	.047	.188	no
Save intention	odds ratio	1.47	$[0.90, 2.41]$	.128	.384	no
Perceived authenticity	mean diff.	0.04	$[-0.08, 0.16]$	.507	1.00	no
Click intention	odds ratio	0.91	$[0.57, 1.47]$	.702	1.00	no

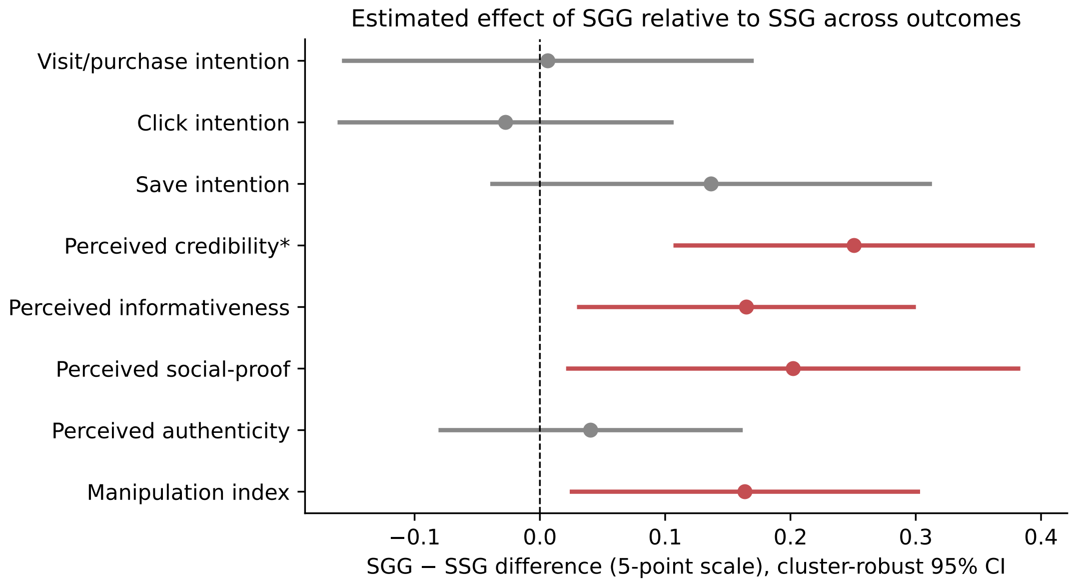

Figure 5: Estimated effect of the SGG pipeline (SGG−SSG) across outcomes, on the common five-point scale, with cluster-robust 95% confidence intervals. Point estimates for the perceptual outcomes favor SGG, but only perceived credibility remains significant after Holm correction; behavioral intentions do not shift.

5.5 Scenario heterogeneity (exploratory)

The null primary effect was uniform across the four scenarios. A model with treatment-by-scenario interactions did not improve fit (joint Wald $p = .94$), and the scenario-specific SGG−SSG differences on visit intention were all small and statistically indistinguishable from zero (ranging from -0.05 to $+0.07$ points; Figure 6). Scenarios did differ substantially in their baseline appeal—mean visit intention ranged from about 3.1 to 3.9 across scenarios—but the treatment moved the outcome in none of them.

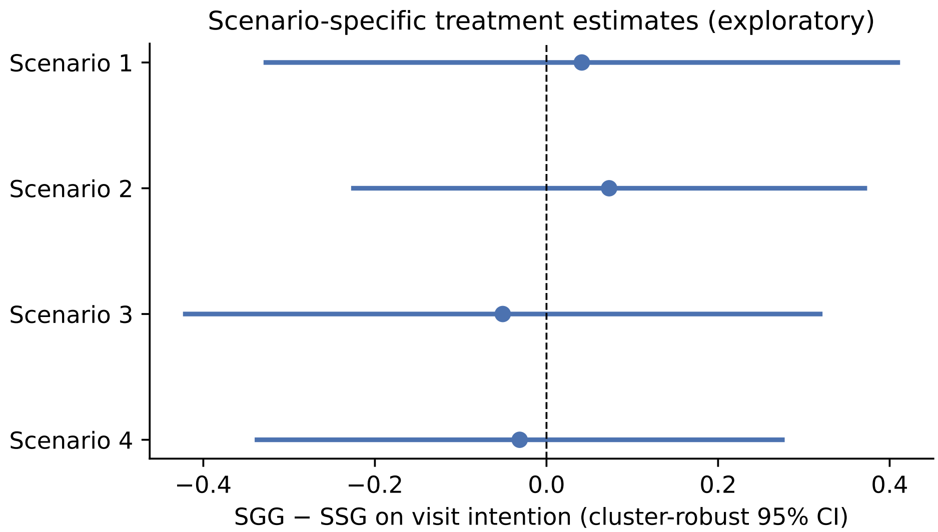

Figure 6: Scenario-specific treatment estimates on visit intention (cluster-robust 95% CIs). The effect is consistently near zero across scenarios.

5.6 Why perceptions did not translate: correlations and mediation-consistency (exploratory)

A natural question is why SGG's gains in perceived credibility, informativeness, and social-proof strength did not raise visit intention. The correlation structure offers a direct answer (Figure 7). Visit intention correlated only weakly with perceived credibility ($r = .01$), informativeness ($r = .08$), social-proof strength ($r = .05$), and authenticity ($r = .05$). In contrast, the perceptual measures correlated moderately *with one another* (e.g., credibility–informativeness $r = .39$, credibility–social-proof $r = .26$), forming a coherent "perceived-quality" cluster that is largely decoupled from stated behavioral intention in this single-exposure setting. Consistent with this, a regression of visit intention on the treatment together with all four perceptions left the (already near-zero) treatment coefficient essentially unchanged, and no perception was a strong predictor of visit intention once the others were included. Because the total treatment effect on visit intention is approximately zero, a formal mediation decomposition is not interpretable here and is not reported; the pattern is presented only as a mechanism-consistent description, in line with cautions about cross-sectional mediation [20, 33, 44].

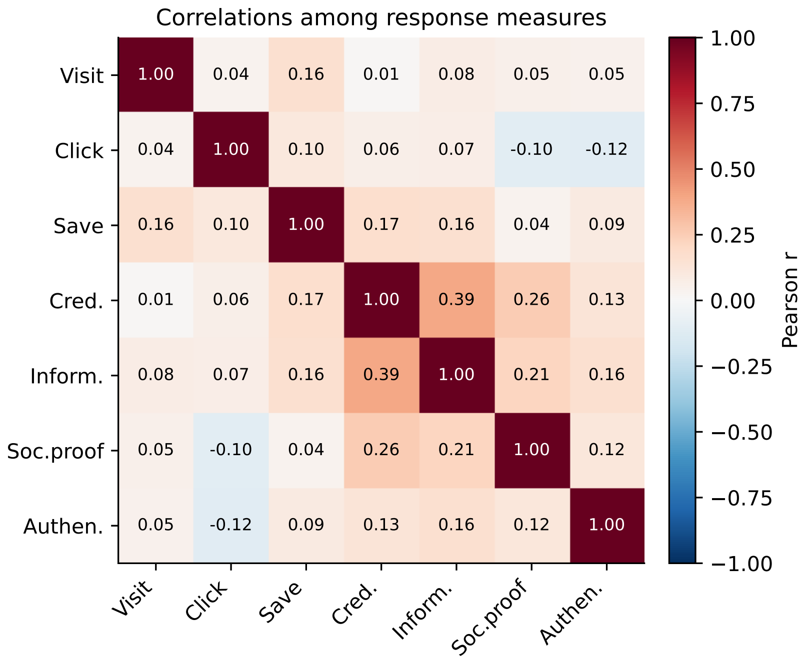

Figure 7: Pearson correlations among response measures. Behavioral intention (visit) is weakly related to the perceptual cluster that the SGG pipeline shifted.

5.7 Exploratory behavioral index and covariate moderation

The preregistered exploratory Behavioral Response Index (a composite of click, save, and visit intention) failed its psychometric pre-check as the three items of the index did not form an internally consistent composite (Cronbach's $\alpha = 0.26$); as specified, this composite was not used to make inferences. Of the exploratory covariate-moderation tests, only Instagram-use frequency showed a borderline interaction with treatment on visit intention ($b = 0.13$, $p = .051$), suggesting that higher social-media use may result in a slightly greater response to SGG, while attitudes toward AI ($p = .66$) and dining-out frequency ($p = .40$) revealed no moderation. These tests are underpowered and exploratory and the results are intended as hypothesis generating only.

5.8 Robustness

The null primary result is robust to specification. A linear model treating visit intention as continuous gave an SGG coefficient of 0.006 points with HC3 robust standard errors ($p = .95$) and an essentially identical estimate with cluster-robust standard errors at the stimulus level (95% CI $[-0.16, 0.17]$, $p = .94$). The agreement between the ordinal and linear specifications, and between model-based and cluster-robust standard errors, indicates that neither the link function nor the clustering assumption drives the conclusion.

5.9 Illustrative pricing exercise

Finally, the results are translated to the illustrative outcome-based pricing exercise, explicitly specified in the design. Let $\widehat{U}$ be the estimated SGG uplift for the high-intent share (those who indicated on visit intention a score of "likely" or "very likely"), and consider the stylized rule

$$PricingSim = F_0 + \lambda \cdot \max(0, \widehat{U}) \cdot \mathbb{1}[\, LB_{95}(\widehat{U}) > 0\,],$$

where F_0 is a fixed base fee, λ a variable multiplier, and the variable component is zero whenever the estimated uplift is negative or statistically inconclusive. In this dataset the high-intent share was 50.4% under SSG and 50.7% under SGG, an uplift of $\widehat{U} = 0.003$ with a stimulus-clustered bootstrap 95% CI of $[-0.13, 0.14]$. Because the lower confidence bound is below zero, the indicator $\mathbb{1}[LB_{95} > 0]$ evaluates to zero and the rule charges only the base fee F_0 (for example, with $F_0 = 100$ and $\lambda = 1000$, the simulated price is 100, with a variable component of 0). The exercise thus illustrates its own guard rail: a transparent outcome-based rule declines to charge for an uplift that cannot be distinguished from noise. This is a conceptual and pedagogical demonstration, not a validated pricing model, and would require field validation against real conversions before any commercial use, particularly given the intention–behavior gap [40, 46].

6 Discussion

The aim of this study was to examine whether Structured Grounded Generation (SGG) leads to a more positive consumer response than Single-Shot Generation (SSG) holding the underlying language model constant and AI authorship undisclosed during exposure. The main result is that the SGG pipeline did not affect the behavioral intention that it was primarily designed to move, but it did affect how the posts were perceived, with judgments of specificity and fact-basedness, and exploratory analyses of credibility, being more positive than under SSG. The implications of this pattern are explored in the theoretical context of Section 2 and presented as consequences in the subsections below, while threats to validity are presented to bound the conclusions.

6.1 Interpreting the perception–behavior dissociation

All three frameworks imply a similar direction of action, and it is not possible to separate their individual actions with this single bundled-treatment design. The dissociation pattern outlined below is not a direct test of the specific claims of ELM, signaling theory, or the MAIN model, but is consistent with each model. SGG was created to increase both the central route (via quality of arguments that are grounded in reality) and the peripheral route (via formatting and credibility cues) according to the Elaboration Likelihood Model. Manipulation checks and the credibility result point to the fact that these cues were indeed received: SGG posts were perceived as more specific, more grounded, and more credible. What did not follow is the downstream step from a more favorable evaluation to a stronger intention to act. In a rapid, one-time exposure, social-media environment, it would seem that the trustworthiness of any given post would have less impact on visit intent than would factors that are comparatively stable, such as appeal of the cuisine or fit of the scenario. This is the same with the data: baseline visit intention was fairly spread out across the four scenarios (roughly 3.1 to 3.9), but the treatment effect was close to zero for each scenario.

Signaling Theory is a compatible argument from a different perspective. The salience of the SGG signals appeared to be adequate to be noticed, but salience is not enough; it raised perceived credibility and the sense that claims were grounded. A signal changes behavior only if it reduces the specific uncertainty that gates the decision. While a consumer's intent to visit is still determined by taste, price, location, or convenience that a single post cannot address, the restaurant's post does become more credible. The MAIN heuristics convey the same message at the level of cues: the authority and bandwagon cues in SGG had an impact on credibility assessments, but not on behavior, indicating that they were working at the evaluative layer rather than the conative layer.

6.2 Why perceptions did not translate

The most obvious evidence on the mechanism is provided by the correlation structure in Section 5.6. There was only a weak correlation between the very perceptions that SGG enhanced (credibility ($r = .01$), informativeness ($r = .08$), and social-proof strength ($r = .05$)) and intention to visit, with perceptions being more highly correlated with each other (r up to .39). The dimensions of "perceived quality" and "stated intention to act" are fairly distinct: in this sample, an intervention that affects the former does not necessarily affect the latter. There are two possibilities to explain this, without excluding each other. First, one exposure might not be enough of an experience to shift a credibility assessment to an intention to visit a restaurant; perceptions change rapidly, but decisions to visit a restaurant accumulate from more than one post. Second, intention in this design could be driven by the presence of a specific preference and scenario fit, thus leaving less room for a content manipulation to account for. Either way, the bottom line is: improvement in content that consistently improves assessment doesn't necessarily affect behavior.

6.3 Implications for outcome-based pricing

This lesson is reinforced by the illustrative pricing exercise of Section 5.9. The uplift estimated with a high intent was not significantly different from zero, so the stylized outcome-based rule only charged its base fee and the variable component of zero. The overall message is cautionary. Assuming a vendor based the price on the perceived credibility or general content quality, they may seem to be providing value that doesn't show up in the behavior which the client paid for. An outcome-based contract should therefore be based not on perceptions but on downstream metrics — clicks, saves, visits or conversions — and should incorporate an explicit statistical safeguard against paying for noise as the indicator term in the pricing rule shows — and should be field-tested, not just based on single-exposure survey intentions. The experiment warns against monetizing a credibility increase, which doesn't propagate to behaviour, in short.

6.4 Relation to prior work and the disclosure boundary

The results align well with research indicating that disclosure is crucial for source- and authorship-based reactions to AI-generated content [4, 22, 24]. Content improvements in the pipeline here shifted perceptions but not behavior, as is thought to be the case when there is a lack of salient AI label, meaning that consumers act on the content, not the origin. The boundary condition of the Persuasion Knowledge Model has not yet been tested by this design, however, and under the new transparency regulations, an explicit 'AI-generated' tag may interact with the quality of the content, in a plausible way, reducing the credibility boost, if not outright eradicating it, given the more structured and polished nature of SGG writing [5, 28, 45]. In the present estimates, we therefore recommend that they be read as bounded to unlabeled-post contexts.

6.5 Limitations and threats to validity

These conclusions were constrained by many limitations. First, all outcomes are one-time, self-reported intentions, and as a result of the intention–behavior gap, a finding on stated intention does not necessarily reflect actual visits or purchases [40, 46]. Second, statistical power is finite, meaning that $N = 261$ is better at detecting modest effects than small ones, but the near unity odds ratio and the near zero linear estimate make it unlikely that a meaningful, but omitted, effect on visit intention specifically, exists. Third, the treatment is a bundled intervention, so without a factorial design, the credibility gain cannot be attributed to any one of the components (structure, grounding, authority cues, or social-proof cues). Moreover, the contrast is one of information as well as process: SSG received only basic restaurant fields, whereas SGG additionally received the authority and social-proof facts. The estimate therefore reflects the full pipeline as deployed, combining structured workflow with richer grounded inputs, rather than structure alone. Fourth, the sample (an online convenience sample of frequent social-media users recruited via student-group channels) and the four restaurant scenarios limited generalizability to other people, product categories, and platforms. Fifth, the preregistered cumulative link mixed model was substituted with its preregistered cluster-robust fallback because for the analysis environment no mixed-ordinal model was available; supplementary diagnostics showed that between-stimulus variance was relatively small, and this fallback is unlikely to have materially influenced the conclusion. Such threats highlight the need for ongoing research as described in Section 7.

7 Conclusions and Further Research

This study examined whether a Structured Grounded Generation (SGG) pipeline, which incorporates structure, factual grounding, embedded authority and social-proof cues, and constraint validation, while using the same base model as the Single-Shot Generation (SSG) pipeline, generates more positive consumer responses to restaurant social-media posts when the AI's authorship is not disclosed. The answer for the confirmatory outcome is no: the SGG pipeline did not increase visit/purchase intention (odds ratio 1.01; 95% CI $[0.66, 1.55]$; $p = .97$), and this null result held across scenarios, and across both ordinal and linear specifications, in a preregistered between-subject experiment with 40 realized stimuli and an analytic sample of 261 participants. Meanwhile, the manipulation was not passive. The posts of the SGG group were perceived as more specific and fact-based; exploratory secondary analyses find that SGG was positively associated with higher perceived credibility (+0.25 points, $p_{\text{Holm}} = .003$; the secondary family is exploratory because the confirmatory test was null). The key empirical finding is the dissociation between these perceptual improvements and behaviour: in this single-exposure, no-disclosure context, there was only a weak correlation between visit intention and the perceptions that the pipeline improved. This illustrative pricing exercise, by design of the exercise, did not charge anything for the variable component since the high-intent uplift was indistinguishable from zero.

These conclusions have several limitations, discussed in Section 6, and suggest specific steps to proceed. Since the outcomes are stated intentions subject to the intention–behavior gap [40, 46], the best follow-up would be a field study that measured click, save and visit behavior, and would place the outcome-based pricing experiment on more solid footing. Since the SGG treatment is a packaged intervention, the next step is to determine which of the treatment's components – structure, grounding, authority cues, or social-proof cues – resulted in the credibility gain. Because the design holds AI authorship undisclosed during exposure, the next logical experiment is a disclosure manipulation where the obligations to disclose contributions of AI that come with the EU AI Act take effect [12, 13, 45]. Lastly, repeated exposure or longitudinal designs could be used to see whether perceptual changes which do not result in a single shot intention do accrue at the brand level over time.

References

[1] Laiba Ali, Faizan Ali, Moh'd Juma Abdalla, and Salman Alotaibi. Beyond the hype: Evaluating the impact of generative AI on brand authenticity, image, and consumer behavior in the restaurant industry. *International Journal of Hospitality Management*, 131:104318, 2025. doi: 10.1016/j.ijhm.2025.104318.

[2] Lars Bergkvist and John R. Rossiter. The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research*, 44(2):175–184, 2007. doi: 10.1509/jmkr.44.2.175.

[3] Godfred O. Boateng, Torsten B. Neilands, Edward A. Frongillo, Heather L. Melgar-Quiñonez, and Sera L. Young. Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6:149, 2018. doi: 10.3389/fpubh.2018.00149.

[4] Jasper David Brüns and Martin Meißner. Do you create your content yourself? using generative artificial intelligence for social media content creation diminishes perceived brand authenticity. *Journal of Retailing and Consumer Services*, 79:103790, 2024. doi: 10.1016/j.jretconser.2024.103790.

[5] Hien Thu Bui. Examining the effect of AI advertising involvement disclosure on advertising value and purchase intentions. *Journal of Research in Interactive Marketing*, 20(5):613–632, 2026. doi: 10.1108/JRIM-02-2025-0066.

[6] A. Colin Cameron and Douglas L. Miller. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372, 2015. doi: 10.3368/jhr.50.2.317.

[7] Qian Chen, Changqin Yin, and Yeming Gong. Would an AI chatbot persuade you: An empirical answer from the elaboration likelihood model. *Information Technology & People*, 38(2):937–962, 2025. doi: 10.1108/ITP-10-2021-0764.

[8] Rune Haubo Bojesen Christensen. *ordinal: Regression Models for Ordinal Data*, 2023. URL https://CRAN.R-project.org/package=ordinal. R package version 2023.12-4.1.

[9] Brian L. Connelly, S. Trevis Certo, R. Duane Ireland, and Christopher R. Reutzel. Signaling theory: A review and assessment. *Journal of Management*, 37(1):39–67, 2011. doi: 10.1177/0149206310388419.

[10] Lee J. Cronbach. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334, 1951.

[11] Adamantios Diamantopoulos, Marko Sarstedt, Christoph Fuchs, Petra Wilczynski, and Sebastian Kaiser. Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *Journal of the Academy of Marketing Science*, 40(3):434–449, 2012. doi: 10.1007/s11747-011-0300-3.

[12] European Commission. Code of practice on marking and labelling of AI-generated content (second draft). https://digital-strategy.ec.europa.eu/en/library/commission-publishes-second-draft-code-practice-marking-and-labelling-ai-generated-content, 2026. Second draft published 5 March 2026; final version pending. European Commission, Shaping Europe's digital future. Transparency obligations under Article 50 of the AI Act become applicable on 2 August 2026.

[13] European Union. Regulation (EU) 2024/1689 of the european parliament and of the council of 13 june 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng, 2024.

[14] Marian Friestad and Peter Wright. The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1):1–31, 1994. doi: 10.1086/209380.

[15] Filippo Gambarota and Gianmarco Altoè. Ordinal regression models made easy: A tutorial on parameter interpretation, data simulation, and power analysis. *International Journal of Psychology*, 59(6):1263–1292, 2024. doi: 10.1002/ijop.13243.

[16] Jamie L. Grigsby, Meg Michelsen, and César Zamudio. Service ads in the era of generative AI: Disclosures, trust, and intangibility. *Journal of Retailing and Consumer Services*, 84:104231, 2025. doi: 10.1016/j.jretconser.2025.104231.

[17] Linda D. Hollebeek, Mark S. Glynn, and Roderick J. Brodie. Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of Interactive Marketing*, 28(2):149–165, 2014. doi: 10.1016/j.intmar.2013.12.002.

[18] Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

[19] Wensheng Huang, Xiaochen Wang, Qianyao Zhang, Junhui Han, and Runtong Zhang. Beyond likes and comments: How social proof influences consumer impulse buying on short-form video platforms. *Journal of Retailing and Consumer Services*, 84:104199, 2025. doi: 10.1016/j.jretconser.2024.104199.

[20] Kosuke Imai, Luke Keele, and Dustin Tingley. A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309–334, 2010. doi: 10.1037/a0020761.

[21] Elvira Ismagilova, Emma L. Slade, Nripendra P. Rana, and Yogesh K. Dwivedi. The effect of electronic word of mouth communications on intention to buy: A meta-analysis. *Information Systems Frontiers*, 22:1203–1226, 2020. doi: 10.1007/s10796-019-09924-y.

[22] Maurice Jakesch, Jeffrey T. Hancock, and Mor Naaman. Human heuristics for AI-generated language are flawed. *Proceedings of the National Academy of Sciences*, 120(11):e2208839120, 2023. doi: 10.1073/pnas.2208839120.

[23] Charles M. Judd, Jacob Westfall, and David A. Kenny. Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1):54–69, 2012. doi: 10.1037/a0028347.

[24] Colleen P. Kirk and Julian Givi. The AI-authorship effect: Understanding authenticity, moral disgust, and consumer responses to AI-generated marketing communications. *Journal of Business Research*, 186:114984, 2025. doi: 10.1016/j.jbusres.2024.114984.

[25] Philip J. Kitchen, Gayle Kerr, Don E. Schultz, Rod McColl, and Heather Pals. The elaboration likelihood model: review, critique and research agenda. *European Journal of Marketing*, 48 (11/12):2033–2050, 2014. doi: 10.1108/EJM-12-2011-0776.

[26] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 2020.

[27] Torrin M. Liddell and John K. Kruschke. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, 79:328–348, 2018. doi: 10.1016/j.jesp.2018.08.009.

[28] Sue Lim and Ralf Schmälzle. The effect of source disclosure on evaluation of AI-generated messages. *Computers in Human Behavior: Artificial Humans*, 2:100058, 2024. doi: 10.1016/j.chbah.2024.100058.

[29] Xialing Lin, Patric R. Spence, and Kenneth A. Lachlan. Social media and credibility indicators: The effect of influence cues. *Computers in Human Behavior*, 63:264–271, 2016. doi: 10.1016/j.chb.2016.05.002.

[30] Chiara Longoni and Luca Cian. Artificial intelligence in utilitarian vs. hedonic contexts: The "word-of-machine" effect. *Journal of Marketing*, 86(1):91–108, 2022. doi: 10.1177/0022242920957347.

[31] Scott B. MacKenzie and Richard J. Lutz. An empirical examination of the structural antecedents of attitude toward the ad in an advertising pretesting context. *Journal of Marketing*, 53(2):48–65, 1989. doi: 10.1177/002224298905300204.

[32] Tamilla Mavlanova, Raquel Benbunan-Fich, and Marios Koufaris. Signaling theory and information asymmetry in online commerce. *Information & Management*, 49(5):240–247, 2012. doi: 10.1016/j.im.2012.05.004.

[33] Scott E. Maxwell and David A. Cole. Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, 12(1):23–44, 2007. doi: 10.1037/1082-989X.12.1.23.

[34] Daniel McNeish and Laura M. Stapleton. Modeling clustered data with very few clusters. *Multivariate Behavioral Research*, 51(4):495–518, 2016. doi: 10.1080/00273171.2016.1167008.

[35] Julie Napoli, Sonia J. Dickinson, Michael B. Beverland, and Francis Farrelly. Measuring consumer-based brand authenticity. *Journal of Business Research*, 67(6):1090–1098, 2014. doi: 10.1016/j.jbusres.2013.06.001.

[36] OECD. Generative AI and the SME workforce. Technical report, OECD Publishing, Paris, 2025. doi: 10.1787/2d08b99d-en.

[37] Roobina Ohanian. Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, 19(3):39–52, 1990. doi: 10.1080/00913367.1990.10673191.

[38] Richard E. Petty and John T. Cacioppo. The elaboration likelihood model of persuasion. In *Communication and Persuasion*, pages 1–24. Springer, New York, 1986. doi: 10.1007/978-1-4612-4964-1_1.

[39] Bruno Schivinski, George Christodoulides, and Dariusz Dabrowski. Measuring consumers' engagement with brand-related social-media content: Development and validation of a scale that identifies levels of social-media engagement with brands. *Journal of Advertising Research*, 56(1):64–80, 2016. doi: 10.2501/JAR-2016-004.

[40] Paschal Sheeran. Intention–behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, 12(1):1–36, 2002. doi: 10.1080/14792772143000003.

[41] Michael Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, 1973. doi: 10.2307/1882010.

[42] S. Shyam Sundar. The MAIN model: A heuristic approach to understanding technology effects on credibility. In Miriam J. Metzger and Andrew J. Flanagin, editors, *Digital Media, Youth, and Credibility*, pages 73–100. MIT Press, Cambridge, MA, 2008. doi: 10.1162/dmal.9780262562324.073.

[43] S. Shyam Sundar, Jinyoung Kim, and Andrew Gambino. Using theory of interactive media effects (TIME) to analyze digital advertising. In Shelly Rodgers and Esther Thorson, editors, *Digital Advertising: Theory and Research*, pages 86–109. Routledge, New York, 2017. doi: 10.4324/9781315623252-6.

[44] Dustin Tingley, Teppei Yamamoto, Kentaro Hirose, Luke Keele, and Kosuke Imai. mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5):1–38, 2014. doi: 10.18637/jss.v059.i05.

[45] Jared Watson, Francesca Valsesia, and Shoshana Segal. Assessing AI receptivity through a persuasion knowledge lens. *Current Opinion in Psychology*, 58:101834, 2024. doi: 10.1016/j.copsyc.2024.101834.

[46] Thomas L. Webb and Paschal Sheeran. Does changing behavioral intentions engender behavior change? a meta-analysis of the experimental evidence. *Psychological Bulletin*, 132(2):249–268, 2006. doi: 10.1037/0033-2909.132.2.249.

[47] Gary L. Wells and Paul D. Windschitl. Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, 25(9):1115–1125, 1999. doi: 10.1177/01461672992512005.

[48] John D. Wells, Joseph S. Valacich, and Traci J. Hess. What signal are you sending? how website quality influences perceptions of product quality and purchase intentions. *MIS Quarterly*, 35(2):373–396, 2011. doi: 10.2307/23044048.

[49] Jacob Westfall, David A. Kenny, and Charles M. Judd. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, 143(5):2020–2045, 2014. doi: 10.1037/xge0000014.