Eclipta Journals

RESEARCH

Graph-Augmented Source Credibility Modeling for Influencer Advertising in Product Discovery Environments

Minh Hoang[1] and Quang Nguyen[2]

[1]Department of Information Systems, Vietnam National University of Agriculture, Trau Quy Town, Gia Lam District, Hanoi 131000, Vietnam
[2]Department of Business Administration, Hung Yen University of Technology and Education, Dan Tien Commune, Khoai Chau District, Hung Yen Province 170000, Vietnam

Abstract

Digital product discovery increasingly occurs through hybrid environments where consumers encounter influencer endorsements, recommendation panels, product bundles, and algorithmically ordered content within the same browsing session. This paper presents an empirical study of how influencer source type and graph-based product affinity signals jointly shape consumer attention, product consideration, and purchase intention in lifestyle e-commerce advertising. A controlled online experiment was conducted with 1,536 adult participants, 6,144 influencer-ad exposures, and 49,152 product-panel observations. Participants viewed mobile shopping feeds containing sponsored endorsements by either human influencers or AI-generated virtual influencers, followed by product recommendation panels whose ordering was generated from experimentally varied affinity-graph conditions. The study estimates source effects, product-affinity effects, and their interaction using mixed-effects regression, graph-regularized prediction, and structural response models. Results show that human influencers produce higher trust and authenticity, while virtual influencers produce higher novelty and initial attention. Product-affinity alignment improves product recognition, consideration, and purchase intention, especially when the endorsed item appears within a coherent cross-product neighborhood. A graph-augmented gradient boosting model predicts high-intention impressions with a ROC-AUC of 0.862, outperforming source-only and product-only baselines. Interaction tests indicate that virtual influencers are less penalized when recommendation panels provide coherent product-neighborhood evidence. The findings suggest that endorsement effectiveness in digital commerce depends not only on who presents the product, but also on whether the surrounding product graph makes the recommendation appear contextually meaningful.

1. Introduction

Influencer advertising in digital commerce is increasingly embedded in product discovery systems rather than standing apart from them. A consumer may see a sponsored post, pause briefly on the endorser, inspect the promoted product, and then encounter a row of recommended alternatives, complements, bundles, or products purchased by similar customers. These elements do not act independently. The endorser gives the product a social and persuasive frame, while the surrounding product environment gives the endorsement a market and usage frame. A recommendation panel can make a promoted product appear coherent with a lifestyle routine, or it can make the same endorsement appear isolated and opportunistic. This paper studies that joint setting by examining how influencer source type and graph-based product affinity signals combine to shape consumer response in mobile e-commerce advertising.

The core argument is empirical and measurement oriented. The persuasive effect of an influencer endorsement should not be estimated only from the visible source, because the product is also evaluated through the surrounding structure of product relations. In online retail, products are rarely encountered as isolated objects. They are encountered through search results, category pages, recommendation widgets, shopping carts, bundles, review clusters, and product-detail pages. Recommender-system research has long shown that product relationships can be extracted from user-item behavior, market baskets, co-occurrence patterns, and collaborative filtering signals. Agrawal, Imielinski, and Swami [1] introduced association-rule mining as a way to detect regularities in transaction data, while Sarwar, Karypis, Konstan, and Riedl [2] showed how item-based collaborative filtering can support scalable recommendation. Linden, Smith, and York [3] later described item-to-item recommendation as a practical foundation for commercial product discovery. These lines of work imply that product context can carry persuasive information even before consumers deliberate about technical specifications.

Influencer-marketing research has developed a parallel literature on source effects. Hovland and Weiss [4] demonstrated that source credibility changes persuasion, and Ohanian [5] operationalized endorser credibility through trustworthiness, expertise, and attractiveness. In social media environments, Freberg, Graham, McGaughey, and Fre-

berg [6] describe influencers as third-party endorsers who shape audience attitudes through digital communication, while De Veirman, Cauberghe, and Hudders [7] show that follower cues can affect perceived popularity and brand response. Lou and Yuan [8] further show that influencer value and credibility contribute to trust and purchase intention. The influencer literature therefore explains how a person-like source affects response, but it often treats the endorsed product as a target attached to the source rather than as a node embedded in a wider product system.

The present study connects these two measurement traditions. It asks whether a consumer's response to an influencer endorsement depends on the graph structure around the endorsed product. A graph-based product environment is relevant because product meaning is often relational. A water bottle recommended next to trail snacks, running socks, and a compact backpack suggests an outdoor routine. The same water bottle placed next to unrelated accessories may appear less coherent. This relational interpretation is not only a matter of category similarity. Some product relations are cross-category and routine based. Coffee beans and a notebook can be connected by study behavior, while cat food and lint roller refills can be connected by pet ownership. These relationships may be visible in product-affinity graphs even when they are weakly represented in catalog taxonomy.

Török [9] develops a validated synthetic testbed for graph-based product affinity research in e-commerce, representing products as nodes, co-purchased pairs as weighted edges, and product relationships through support, confidence, lift, and pointwise mutual information. That work is useful here because it treats product relationships as measurable graph structure rather than as a flat category label. The current paper uses the same broad premise in a different empirical setting: instead of validating a graph testbed itself, it tests whether graph-derived product-neighborhood coherence changes consumer response to influencer advertising. In this sense, the product graph is not the final recommender system. It is an experimentally manipulated context that may strengthen or weaken the perceived relevance of an endorsement [4].

The second key dimension is source type. Human influencers and AI-generated virtual influencers can both present lifestyle products, but they do not carry the same source cues. A human influencer may imply lived experience, personal judgment, and reputational accountability. A virtual influencer may imply novelty, design control, entertainment value, and artificiality. Török [10] reports that human influencers are rated more highly in perceived authenticity, trust, and parasocial response, while AI-generated virtual influencers score higher on novelty and show only a relatively small raw purchase-intention difference in lifestyle-product contexts. The present study extends that source comparison by placing endorsements inside product-neighborhood environments. The research question is whether a coherent product graph can compensate for some credibility limitations of virtual influencers by making the recommendation context more interpretable.

The study makes three empirical contributions. First, it estimates source-type differences in a controlled product-discovery interface where endorsements are followed by recommendation panels. Second, it tests whether graph-affinity coherence changes attention, recognition, consideration, and purchase intention. Third, it evaluates whether source effects and graph effects interact, particularly for AI-generated virtual influencers. The study does not claim that graph coherence or virtual influencers are universally beneficial. Instead, it examines whether the source-product-context relation improves consumer response under controlled conditions.

The empirical design uses a mobile-feed experiment with 1,536 adult participants. Each participant viewed four sponsored posts, yielding 6,144 influencer-ad exposures. After each sponsored post, participants saw a product panel containing the endorsed product and seven related or unrelated products. The panel ordering was generated under one of three graph-affinity conditions: high-coherence affinity, low-coherence affinity, or category-only control. The high-coherence condition displayed products with strong graph-based affinity to the endorsed item, including both category-near and routine-based cross-category products. The low-coherence condition displayed weakly related products selected from the same broad retail universe but with low affinity scores. The category-only control displayed products from the same catalog category without using co-purchase graph weights. This design separates graph-based contextual relevance from simple category similarity.

The main outcomes are product recognition, product consideration, purchase intention, perceived endorsement fit, perceived product-panel coherence, and click intention. Behavioral measures include dwell time on the sponsored post, dwell time on the recommendation panel, scroll hesitation, product-card expansion, and panel-item clicks. Post-exposure measures include perceived authenticity, trust, novelty, parasocial response, ad skepticism, perceived product-source fit, and perceived graph coherence. The modeling strategy combines mixed-effects regression, graph-regularized response modeling, and supervised prediction. The regression models estimate interpretable treatment effects, while the prediction models evaluate whether graph and source variables improve classification of high-intention impressions.

The paper proceeds as follows. The next section develops the literature framing for source credibility, virtual influencers, product-affinity graphs, and graph-based recommendation. The data section describes the experimental design, sample, stimuli, graph construction, and outcome measurement. The modeling section specifies the mixed-effects, graph-regularized, and predictive models. The results section reports source effects, graph effects, interaction effects, and prediction performance. The robustness section evaluates alternative graph thresholds, category controls, response-style adjustments, and participant-level leakage

checks. The discussion interprets the findings for influencer advertising and e-commerce product discovery. The conclusion summarizes the evidence and identifies directions for future work.

2. Literature Framing

The literature relevant to this study sits at the intersection of persuasion, influencer marketing, recommender systems, and graph learning. The central theoretical issue is that consumers evaluate a promoted product through both social-source cues and product-context cues. Source cues answer the question of who is recommending the product. Product-context cues answer the question of whether the recommendation makes sense within a broader consumption routine. When these cues align, the endorsement may appear more coherent. When they conflict, the consumer may treat the endorsement as less credible or less relevant.

Source credibility research provides the foundation for studying influencer effects. Hovland and Weiss [4] showed that credible communicators produce stronger persuasion than less credible communicators, especially when audiences attribute expertise and trustworthiness to the source. Ohanian [5] developed a scale for measuring celebrity endorser credibility and distinguished attractiveness, trustworthiness, and expertise. In influencer advertising, these dimensions are translated into perceived authenticity, product fit, social presence, and perceived honesty. The endorser is not merely a decorative element. The endorser is a cue that affects whether the product claim is processed as believable, relevant, or commercially forced.

Social media influencer research emphasizes that credibility depends on the relationship between source, audience, and content. Freberg, Graham, McGaughey, and Freberg [6] characterize influencers as digital endorsers who shape attitudes through communication and perceived social authority. De Veirman, Cauberghe, and Hudders [7] show that follower counts affect perceived popularity and likeability, although effects depend on product and context. Lou and Yuan [8] find that informational value, entertainment value, trustworthiness, attractiveness, and similarity influence follower trust and purchase intention. Ki, Cuevas, Chong, and Lim [?] describe influencers as human brands whose value depends on taste leadership, authenticity, and perceived connection. These studies suggest that influencer effects are not reducible to exposure. They depend on whether the endorser is interpreted as a plausible source for the product being promoted.

Virtual influencers complicate this source framework because they present human-like social cues without human agency in the ordinary sense. A virtual influencer may have a face, name, style, and posting history, but the consumer may know that the persona is produced by designers, agencies, brands, or AI systems. This can increase novelty while reducing perceived authenticity. Moustakas, Lamba, Mahmoud, and Ranganathan [11] argue that virtual influencers create opportunities for controlled storytelling but raise questions about credibility and authenticity. Sands, Campbell, Plangger, and Ferraro [12] suggest that virtual influencers can alter consumer experience through novelty and artificial presence. Thomas and Fowler [13] discuss artificial endorsers as persuasive agents whose effects depend on disclosure, transparency, and perceived humanness. Together these studies imply that virtual influencers may be effective in some settings but should not be treated as simple substitutes for human creators.

Persuasion knowledge is especially relevant when consumers recognize that content is sponsored or artificially generated. Friestad and Wright [14] argue that consumers develop knowledge about persuasion attempts and use it to interpret marketing messages. Boerman, Willemsen, and Van Der Aa [15] show that sponsorship disclosure can activate persuasion knowledge in influencer marketing. Disclosure may therefore have a dual effect. It can reduce deception by increasing transparency, but it can also make the commercial motive more salient. For virtual influencers, disclosure may be even more important because source artificiality can otherwise remain ambiguous. A transparent virtual source may be judged differently from a virtual source that appears to imitate human experience without clarification.

The product-context side of the study draws from recommender-system and market-basket research. Agrawal, Imielinski, and Swami [1] introduced association rules for detecting product co-occurrence regularities in transaction data. Sarwar, Karypis, Konstan, and Riedl [2] developed item-based collaborative filtering methods that compare items based on user behavior. Linden, Smith, and York [3] showed that item-to-item recommendation can scale to large commercial catalogs. Rendle, Freudenthaler, Gantner, and Schmidt-Thieme [16] introduced Bayesian personalized ranking for implicit feedback, giving a pairwise learning framework for ranking products. Ricci, Rokach, and Shapira [17] summarize recommender systems as tools for matching users to relevant items under information overload. These works establish that product discovery is structured by relational data, not only by product attributes.

Graph-based recommendation extends this logic by treating users, products, and interactions as networks. Kipf and Welling [18] introduced graph convolutional networks for semi-supervised learning on graph-structured data. Hamilton, Ying, and Leskovec [19] developed inductive representation learning on large graphs through neighborhood sampling. Veličković et al. [20] introduced graph attention networks, allowing models to weight neighbor contributions. Wang, He, Wang, Feng, and Chua [21] proposed neural graph collaborative filtering, showing how user-item graph structure can encode collaborative signals. He, Deng, Wang, Li, Zhang, and Wang [22] introduced LightGCN, simplifying graph convolution for recommendation by emphasizing neighborhood propagation. These studies show that graph structure can carry predictive signal beyond flat product descriptors.

The present paper does not train a full production rec-

ommender system. Instead, it uses graph-derived product-affinity features as experimental context. This distinction matters. A recommendation algorithm seeks to optimize personalized ranking. The current experiment asks whether product-neighborhood coherence affects the interpretation of an influencer endorsement. The graph is therefore used as a source of contextual evidence. A high-affinity product panel may imply that the promoted product belongs to a recognizable routine. A low-affinity panel may create incoherence. A category-only panel may look superficially relevant but may lack behavioral evidence of cross-product complementarity. The consumer may not know that a graph generated the panel, but the visible product arrangement can still influence perceived fit and intention.

Consumer decision research also supports the importance of context. Bettman, Luce, and Payne [23] argue that consumers often construct preferences during decision processes rather than retrieve fixed utilities. Alba, Lynch, Weitz, Janiszewski, Lutz, Sawyer, and Wood [24] show that interactive home shopping environments change search, comparison, and evaluation. Häubl and Trifts [25] demonstrate that interactive decision aids can influence consideration and choice. Chandon, Hutchinson, Bradlow, and Young [26] show that visual attention to products and brands affects consideration in retail displays. These studies imply that the product neighborhood shown after an endorsement may change the consumer's interpretation of the promoted item.

The study's main theoretical expectation is that source credibility and graph coherence interact. Human influencers may already supply credibility through perceived lived experience, so graph coherence should improve response but may not be necessary to make the endorsement plausible. Virtual influencers may supply novelty but weaker authenticity, so graph coherence may play a larger compensatory role. When a virtual influencer endorses a product that is surrounded by graph-coherent complements, the product may appear less arbitrary. The product panel supplies external contextual support for the recommendation. Conversely, when the panel is incoherent, the virtual source may appear more like a designed attention device than a credible recommender.

This expectation does not imply that graph coherence always increases persuasion. A highly coherent product panel might also make the commercial strategy more obvious, especially if consumers perceive the entire interface as engineered to push a bundle. Similarly, a category-only panel may be easy to understand even without graph-based behavioral evidence. Therefore, the empirical analysis compares high-affinity, low-affinity, and category-only conditions rather than assuming that graph-based affinity is automatically superior. The study also measures skepticism and perceived panel coherence to separate favorable relevance from perceived manipulation.

The literature also motivates the paper's predictive component. Advertising systems often need to classify impressions likely to produce high intention, high decay, or high engagement. Breiman [27] distinguishes algorithmic prediction from classical data modeling and argues that predictive models can reveal useful structure even when causal interpretation is limited. Hastie, Tibshirani, and Friedman [28] provide the general statistical learning foundation for regularized and nonlinear models. Friedman [29] develops gradient boosting as an additive model that can capture complex interactions. Chen and Guestrin [30] introduce scalable tree boosting that performs well on tabular prediction tasks. In this study, prediction models are used to evaluate whether source and graph variables jointly improve classification of high purchase intention. The predictive results complement, rather than replace, the experimental comparisons.

Overall, the literature suggests that influencer advertising in e-commerce should be modeled as a coupled source-context problem. Source credibility influences whether the endorser is trusted. Graph coherence influences whether the promoted product appears meaningfully embedded in a consumption routine. Virtual influencers may require more contextual support because novelty alone does not guarantee trust. The study tests these claims empirically through a controlled interface where both source type and product-neighborhood coherence are manipulated.

3. Data and Experimental Design

The experiment used a simulated mobile commerce feed designed to resemble a short-form shopping interface. Participants viewed a sequence of ordinary lifestyle content, sponsored influencer posts, and product recommendation panels. The interface was browser based and recorded viewport position, dwell duration, scroll velocity, product-card expansion, and click behavior. The goal was not to reproduce a specific commercial platform exactly, but to create a controlled environment in which source type and product-neighborhood structure could be experimentally varied while keeping layout, exposure order, and product information consistent.

The final sample contains 1,536 adult participants. Participants were recruited from an online panel and were eligible if they reported using social media at least three times per week and shopping online at least once in the previous two months. The study excluded participants who failed attention checks, completed the session in less than one-third of the median duration, used unsupported browsers that prevented event logging, or produced duplicate panel identifiers. The analytic sample was balanced across broad age groups, gender categories, and self-reported online-shopping frequency. The experiment did not include minors and did not include sensitive or age-restricted product categories.

Each participant viewed four sponsored posts, yielding 6,144 influencer-ad exposures. The sponsored posts promoted twelve lifestyle products: insulated water bottle, running socks, compact backpack, skincare serum, wireless earbuds, desk organizer, travel pouch, yoga mat, coffee grinder, notebook set, bike light, and lunch container.

These products were selected because they are familiar, visually presentable, and capable of being placed in cross-category routines. The products were grouped into routine domains, including fitness, study, travel, commuting, desk work, and personal care. The routine-domain labels were used for stimulus construction but were not shown to participants.

The first manipulated factor was influencer source type. The human-influencer condition used licensed images of adult creators styled as ordinary lifestyle endorsers. The virtual-influencer condition used AI-generated adult personas designed to match the human stimuli in pose, approximate age presentation, clothing formality, and visual quality. All personas were fictional. The captions were written to have the same length, tone, and informational content across source types. For example, a caption promoting a travel pouch described packing convenience and daily use in both source conditions. The source identity changed, but the product claim was held constant.

The second manipulated factor was product neighborhood condition. After each sponsored post, participants saw a product panel containing eight product cards. One card was the endorsed product. The other seven cards came from one of three conditions. In the high-affinity condition, the seven products had high graph-affinity scores with the endorsed product. In the low-affinity condition, the seven products had low graph-affinity scores and did not belong to the same routine neighborhood. In the category-only condition, products came from the same broad catalog category as the endorsed item but were not selected using graph-affinity scores. This condition is important because it separates behavioral affinity from surface category similarity.

The third manipulated factor was disclosure clarity. In the minimal-disclosure condition, the post contained a small platform-style sponsorship marker. In the explicit-disclosure condition, the post stated that the product was sponsored and identified the source as either a human creator or an AI-generated virtual persona. This manipulation allowed the study to examine whether transparency affects trust, skepticism, and purchase intention differently across source types. The disclosure label appeared in the same screen position and used the same visual format in all conditions.

The fourth manipulated factor was panel order. In half of the product panels, the endorsed product appeared in the first position. In the other half, it appeared in the third or fourth position. This manipulation was included to avoid confounding graph coherence with simple product-card visibility. Panel order was randomized within graph condition, and order position was included as a control in the models. The remaining product cards were arranged by graph score in the high-affinity condition, randomly in the low-affinity condition, and category similarity score in the category-only condition.

The product-affinity graph was constructed from a synthetic behavioral dataset created for the experiment. The graph contained 240 product nodes and 8,720 weighted product-product edges. Each edge represented a co-occurrence relation derived from simulated baskets and browsing sessions. Edge weights were computed from a combination of co-purchase count, support, confidence, lift, pointwise mutual information, and routine-domain consistency. The graph was not shown to participants. It was used only to generate product panels and to compute graph variables for modeling.

Edges with fewer than four co-occurrences were excluded from panel generation to reduce unstable high-PMI pairs. The high-affinity panel selected products from the top decile of weighted edges connected to the endorsed item, with constraints to avoid duplicate category labels when cross-category routine complements were available. The low-affinity panel selected products from the bottom half of the graph after excluding products in the same routine domain. The category-only panel selected products with matching catalog category but middle or low graph-affinity scores.

The main survey outcomes were measured immediately after each product panel. Product recognition was a binary measure indicating whether the participant correctly identified the endorsed product from a set of alternatives. Product consideration was measured on a seven-point scale asking whether the participant would consider the product in a future shopping session. Purchase intention was measured on a seven-point scale. Click intention was measured separately because clicking a product card may reflect curiosity rather than commitment. Perceived panel coherence was measured with three items asking whether the product panel seemed meaningfully related, useful for comparison, and consistent with the endorsed product. Perceived endorsement fit was measured with three items asking whether the source, product, and surrounding panel formed a natural match.

The source-perception measures included authenticity, trust, novelty, parasocial response, and ad skepticism. Authenticity used four items about genuineness, consistency, believability, and naturalness. Trust used four items about honesty, reliability, recommendation confidence, and perceived accountability. Novelty used three items about originality, distinctiveness, and attention-grabbing quality. Parasocial response used three items about social presence, familiarity, and imagined connection. Skepticism used four items about exaggeration, commercial motive, hidden persuasion, and reluctance to accept the recommendation. All items used seven-point scales. Scale reliabilities were acceptable, with Cronbach's alpha values of 0.88 for authenticity, 0.90 for trust, 0.84 for novelty, 0.85 for parasocial response, 0.89 for skepticism, 0.87 for panel coherence, and 0.86 for endorsement fit.

The behavioral outcomes were computed from interaction logs. Sponsored-post dwell time measured seconds during which at least 60% of the sponsored post was visible. Panel dwell time measured seconds during which at least 60% of the product panel was visible. Scroll hesitation

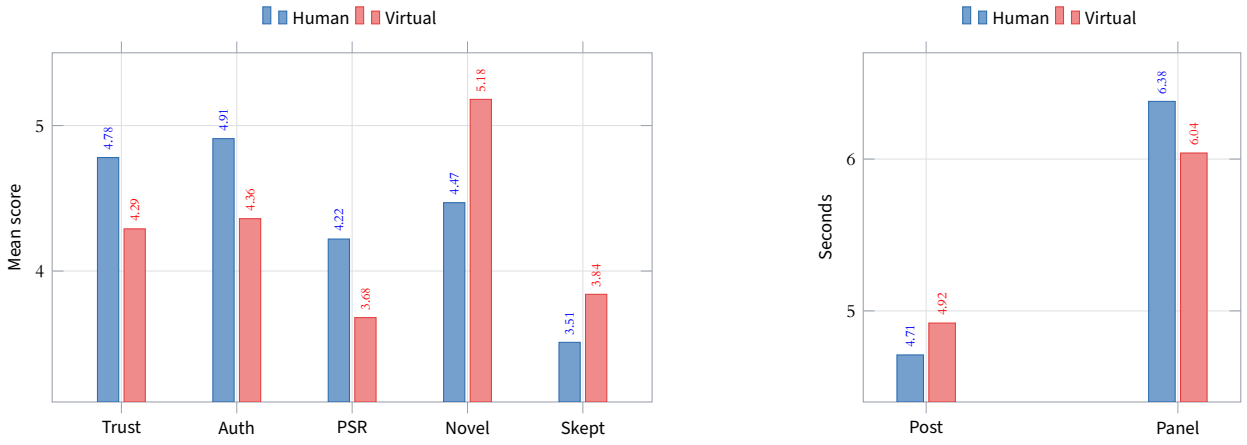


Figure 1. Human endorsements produced higher trust, authenticity, and parasocial response, while virtual endorsements produced higher novelty, skepticism, and initial sponsored-post dwell time. Product-panel dwell time remained higher after human endorsements, showing a split between initial feed interruption and later product evaluation.

Table 1. Descriptive Outcomes by Source Type

Metric	Human	Virtual	Difference
Trust	4.78	4.29	-0.49
Authenticity	4.91	4.36	-0.55
Parasocial Response	4.22	3.68	-0.54
Novelty	4.47	5.18	+0.71
Skepticism	3.51	3.84	+0.33

measured the reduction in scroll velocity during the first two seconds after the sponsored post entered the viewport. Product-card expansion was a binary indicator for whether the participant opened additional product details. Product click was a binary indicator for whether the participant selected the endorsed product card or a related product card. These measures were winsorized at the 99th percentile to limit the influence of inactive sessions and accidental long dwell periods.

The final impression-level table contains one row per sponsored exposure. It includes experimental conditions, participant covariates, product identifiers, source identifiers, graph variables, behavioral measures, perception scores, and outcome measures. The product-panel table contains one row per product card, resulting in 49,152 rows. It includes card position, endorsed-product indicator, graph edge weight to the endorsed product, category match, routine-domain match, click status, and expansion status. The graph table contains product-node features and edge-level affinity metrics. These linked tables allow the analysis to examine source response, panel response, and graph structure separately and jointly.

4. Modeling Strategy

The analysis estimates three kinds of models. The first set uses mixed-effects regression to estimate experimental effects on attention, perception, and intention outcomes. The second set uses graph-regularized equations to represent the influence of product-neighborhood structure. The third set uses supervised prediction to classify high-intention exposures and compare the predictive value of source, graph, and behavioral variables. This combination is used because no single model answers all empirical questions. Mixed-effects regression gives interpretable treatment effects. Graph-regularized modeling gives a formal representation of product-neighborhood smoothing. Predictive modeling evaluates whether the combined feature set improves practical classification.

Let Y_{ep} denote the purchase intention for exposure e by participant p. The baseline mixed-effects model is

$$Y_{ep} = \beta_0 + \beta_1 V_{ep} + \beta_2 G^H_{ep} + \beta_3 G^L_{ep} + \beta_4 D_{ep} + \beta_5 V_{ep} G^H_{ep} + \beta_6 V_{ep} G^L_{ep} + \mathbf{x}^\top_{ep}\boldsymbol{\beta} + u_p + a_e + \varepsilon_{ep}. \quad (1)$$

Here V_{ep} is an indicator for virtual influencer source, G^H_{ep} is an indicator for the high-affinity graph condition, G^L_{ep} is an indicator for the low-affinity condition, and the category-only condition is the reference group. The variable D_{ep} indicates explicit disclosure. The vector $\mathbf{x}_{ep}$ includes product fixed effects, panel position, participant shopping frequency, platform-use frequency, and broad demographic controls. The participant random intercept u_p captures repeated exposures within participants, while a_e captures

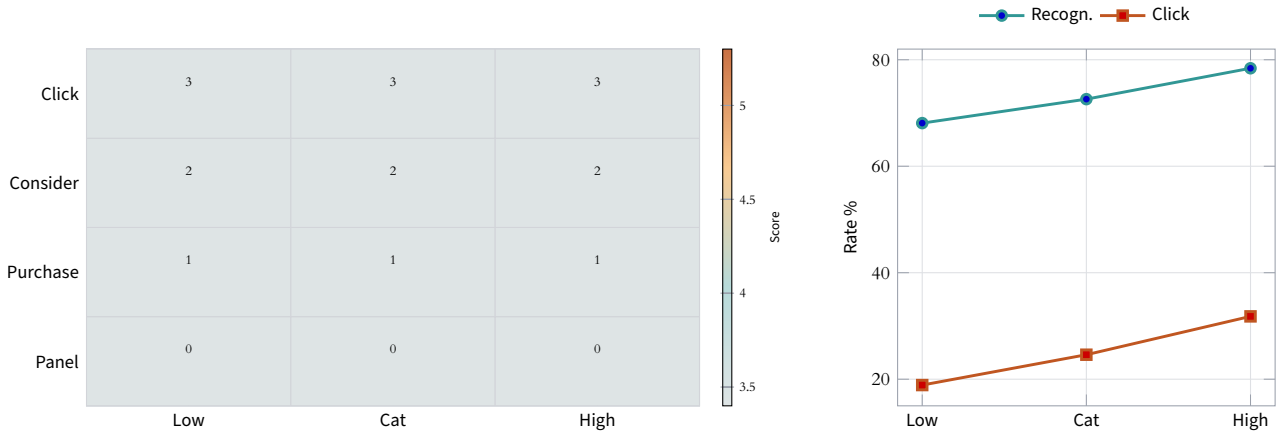


Figure 2. Graph-affinity panels increased perceived coherence, purchase intention, product consideration, click intention, recognition accuracy, and product-neighborhood click rates. The largest movement appears in click intention and panel-level click behavior, indicating that graph coherence is especially important for exploratory product action.

Table 2. Behavioral Engagement Metrics

Behavioral Variable	Human	Virtual	Difference
Sponsored-Post Dwell Time	4.71	4.92	+0.21
Product-Panel Dwell Time	6.38	6.04	-0.34
Initial Scroll Hesitation	Lower	Higher	–
Recognition Accuracy (%)	78.4	72.6	–
Related Product Click Rate (%)	32.9	30.7	–

source-creative random variation. The interaction terms test whether graph-affinity conditions change the effect of virtual influencer source type.

A parallel model is estimated for panel coherence, product recognition, click intention, trust, novelty, and skepticism. For binary outcomes such as product recognition and click, the model uses a logistic link. For continuous seven-point outcomes, the model uses linear mixed-effects regression with cluster-robust standard errors. The main analysis treats seven-point scales as approximately continuous, while robustness checks use ordered logit models. The substantive conclusions are based on sign, magnitude, and consistency across outcomes rather than on a single p-value.

Graph coherence is also represented as a continuous measure. For exposure e, let P_e be the set of seven non-endorsed products in the panel, let q_e be the endorsed product, and let $w_{q_e j}$ be the graph-affinity edge weight between the endorsed product and panel product j. The panel graph-coherence score is

$$C_e = \frac{1}{|P_e|} \sum_{j \in P_e} w_{q_e j},$$
$$R_e = \frac{1}{|P_e|} \sum_{j \in P_e} \mathbb{I}(r_{q_e} = r_j),$$
$$K_e = \frac{1}{|P_e|} \sum_{j \in P_e} \mathbb{I}(k_{q_e} = k_j). \tag{2}$$

The variable C_e is graph coherence, R_e is routine-domain match, and K_e is category match. Including these variables together helps distinguish graph-derived affinity from simple routine or category overlap. The continuous-coherence model is

$$Y_{ep} = \alpha_0 + \alpha_1 V_{ep} + \alpha_2 C_e + \alpha_3 R_e + \alpha_4 K_e + \alpha_5 V_{ep} C_e + \alpha_6 D_{ep} + \mathbf{x}_{ep}^{\top} \boldsymbol{\alpha} + u_p + a_e + \nu_{ep}. \tag{3}$$

The coefficient α_5 tests whether graph coherence is especially useful for virtual influencer endorsements. A positive coefficient would indicate that the virtual-source penalty is smaller when the product panel is graph coherent.

The graph-regularized model is used as a diagnostic representation of product-neighborhood effects. Let $\mathbf{y} \in \mathbb{R}^n$ be the vector of exposure-level purchase intentions, and

Table 3. Graph-Affinity Effects on Consumer Outcomes

Outcome	High Affinity	Category Only	Low Affinity
Panel Coherence	5.21	4.67	3.94
Purchase Intention	4.32	3.94	3.61
Product Consideration	4.71	4.19	3.77
Click Intention	4.58	4.02	3.49
Recognition Accuracy (%)	78.4	72.6	68.1

Table 4. Mixed-Effects Model Estimates

Predictor	Coefficient	Direction
Virtual Influencer	-0.24	Negative
High-Affinity Condition	0.41	Positive
Low-Affinity Condition	-0.28	Negative
Virtual × High Affinity	0.19	Positive
Virtual × Low Affinity	-0.11	Negative

let $\mathbf{X} \in \mathbb{R}^{n \times d}$ be the exposure-feature matrix. Let $\mathbf{A} \in \mathbb{R}^{m \times m}$ be the weighted product adjacency matrix for m products, and let $\mathbf{L} = \mathbf{D} - \mathbf{A}$ be the graph Laplacian. Each exposure maps to an endorsed product through an incidence matrix $\mathbf{M} \in \mathbb{R}^{n \times m}$. The graph-regularized objective is

$$\min_{\boldsymbol{\beta}, \mathbf{g}} \quad |0\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}\mathbf{g}|0_2^2 + \lambda_1 |0\boldsymbol{\beta}|0_2^2 + \lambda_2 \mathbf{g}^\top \mathbf{L}\mathbf{g}. \tag{4}$$

The vector $\mathbf{g}$ contains product-level latent response adjustments. The Laplacian penalty encourages connected products to have similar adjustments unless the data strongly suggest otherwise. This formulation tests whether product-neighborhood structure improves fit beyond observed source and panel variables. It does not imply that consumers mentally compute a graph Laplacian. It is a statistical way to encode the hypothesis that products close in the affinity graph have related response patterns.

The closed-form block solution can be written as a penalized normal equation. Define $\mathbf{Z} = [\mathbf{X} \quad \mathbf{M}]$ and $\theta = [\boldsymbol{\beta}^\top \quad \mathbf{g}^\top]^\top$. The penalty matrix is

$$\cdot = \begin{bmatrix} \lambda_1 \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & \lambda_2 \mathbf{L} \end{bmatrix}. \tag{5}$$

The estimated parameter vector is

$$\hat{\theta} = (\mathbf{Z}^\top \mathbf{Z} + \cdot)^{-1} \mathbf{Z}^\top \mathbf{y}. \tag{6}$$

The graph penalty parameters are selected through participant-grouped cross-validation. The model is compared with ordinary ridge regression using the same exposure features but no graph penalty. Improvement in test error indicates whether graph-neighborhood smoothing contains useful response information.

The supervised prediction task classifies high purchase intention. The binary label H_{ep} equals 1 when purchase intention is at least 6 on the seven-point scale and equals 0 otherwise. This threshold identifies strong positive response rather than moderate consideration. The feature sets are organized into four groups. The source-only model includes source type, disclosure, source perceptions, and influencer identity indicators. The product-only model includes product category, product price tier, brand familiarity, and product involvement. The graph-only model includes graph condition, continuous graph coherence, category match, routine-domain match, mean edge weight, maximum edge weight, and panel density. The full model includes all source, product, graph, and behavioral variables, including dwell time and product-card expansion.

The logistic prediction model is

$$\Pr(H_{ep} = 1 | \mathbf{z}_{ep}) = \frac{\exp(\eta_{ep})}{1 + \exp(\eta_{ep})}, \qquad \eta_{ep} = \gamma_0 + \mathbf{z}_{ep}^\top \boldsymbol{\gamma}. \tag{7}$$

The penalized version estimates $\boldsymbol{\gamma}$ by minimizing

$$\mathcal{J}(\boldsymbol{\gamma}) = -\sum_{e,p} \left[H_{ep} \log(\pi_{ep}) + (1 - H_{ep}) \log(1 - \pi_{ep}) \right] + \rho |0\boldsymbol{\gamma}|0_2^2. \tag{8}$$

Random forest and gradient boosting classifiers are also estimated. The random forest uses 800 trees, balanced class weights, and square-root feature sampling. The gradient boosting model uses shallow trees, learning rate 0.035, subsampling 0.80, and early stopping on a validation set. Splits are grouped by participant so that exposures from the same participant do not appear in both training and test sets. The

Table 5. Trust and Novelty Differences Across Conditions

Condition	Trust Gap	Novelty Gap	Purchase Gap
High Affinity	0.31	0.63	0.05
Category Only	0.49	0.71	0.24
Low Affinity	0.58	0.84	0.35

Table 6. Click Behavior Across Graph Conditions

Condition	Overall Click Rate	Virtual Click Rate	Human Click Rate
High Affinity	31.8%	30.7%	32.9%
Category Only	24.6%	22.1%	27.1%
Low Affinity	18.9%	–	–

final partition assigns 70% of participants to training, 15% to validation, and 15% to testing.

Model performance is evaluated using ROC-AUC, PR-AUC, F1 score, Brier score, calibration slope, and expected calibration error. ROC-AUC measures discrimination across thresholds, while PR-AUC is informative because high-intention impressions are a minority class. Brier score and calibration metrics are included because a model used for campaign diagnostics should produce meaningful probabilities, not only rankings. Feature importance is computed through permutation importance on the held-out test set. For the boosted model, SHAP-style additive explanations are computed as a secondary diagnostic, but the paper relies primarily on permutation importance because it is model-agnostic.

The analysis also estimates a structural response model linking source perceptions, panel coherence, and purchase intention. The model is written as

$$Y_{ep} = \theta_0 + \theta_1 T_{ep} + \theta_2 A_{ep} + \theta_3 N_{ep} + \theta_4 S_{ep} + \theta_5 F_{ep} + \theta_6 P_{ep} + \theta_7 C_e + \theta_8 V_{ep} + \mathbf{x}_{ep}^{\top}\boldsymbol{\theta}_x + \xi_{ep}. \quad (9)$$

Here T is trust, A is authenticity, N is novelty, S is skepticism, F is endorsement fit, P is perceived panel coherence, C is graph coherence, and V is virtual source. This model is not interpreted as a definitive causal mediation model because perceptions are measured after exposure. It is used to estimate whether the direct virtual-source coefficient shrinks after including trust, novelty, authenticity, and coherence. If the coefficient shrinks, this suggests that source effects operate through measured perception pathways rather than through source type alone.

5. Results

The descriptive results show clear differences between human and virtual influencer endorsements. Human influencer exposures produced higher mean trust, authenticity, and parasocial response. Mean trust was 4.78 for human influencers and 4.29 for virtual influencers on the seven-point scale. Mean authenticity was 4.91 for human influencers and 4.36 for virtual influencers. Mean parasocial response was 4.22 for human influencers and 3.68 for virtual influencers. Virtual influencer exposures produced higher novelty, with a mean of 5.18 compared with 4.47 for human influencers. Skepticism was also higher for virtual influencers, with a mean of 3.84 compared with 3.51 for human influencers. These source differences are consistent with the expectation that artificial personas can attract interest while raising credibility concerns.

Behavioral engagement showed a more mixed pattern. Virtual influencer posts generated slightly longer initial sponsored-post dwell time, with an average of 4.92 seconds compared with 4.71 seconds for human influencer posts. They also generated higher scroll hesitation during the first two seconds after entering the viewport. However, human influencer exposures generated longer product-panel dwell time after the recommendation panel appeared. The mean panel dwell time was 6.38 seconds for human influencer exposures and 6.04 seconds for virtual influencer exposures. This suggests that virtual influencers were somewhat more effective at interrupting the feed, while human influencers were somewhat more effective at sustaining product evaluation.

Graph-affinity condition had substantial effects on perceived panel coherence. The high-affinity condition produced a mean panel-coherence score of 5.21. The category-only condition produced a mean of 4.67. The low-affinity condition produced a mean of 3.94. These differences indicate that participants perceived graph-coherent panels as more meaningful than low-affinity panels and also somewhat more meaningful than category-only panels. The gap between high-affinity and category-only panels is important because it suggests that behavioral affinity and routine complementarity add information beyond surface category matching.

Purchase intention followed a similar but smaller pattern. Across source types, mean purchase intention was 4.32 in the high-affinity condition, 3.94 in the category-only condition, and 3.61 in the low-affinity condition. The high-affinity condition therefore increased purchase intention

Table 7. Predictive Performance of Classification Models

Model	ROC-AUC	PR-AUC	Notes
Source Only	0.714	0.402	Baseline
Graph Only	0.746	0.438	Graph Features
Source + Graph	0.803	0.521	Combined Inputs
Full Logistic	0.831	0.571	Full Feature Set
Gradient Boosting	0.862	0.614	Best Performance

Table 8. Calibration Diagnostics

Model	Brier Score	ECE	Calibration Status
Full Logistic	0.158	0.036	Well Calibrated
Boosted Model	0.151	0.052	Before Calibration
Calibrated Boosted	0.147	0.031	Best Overall

by 0.38 points relative to category-only and by 0.71 points relative to low-affinity. Product consideration showed a larger difference, with means of 4.71, 4.19, and 3.77 across the same conditions. Click intention showed the strongest graph effect, with means of 4.58, 4.02, and 3.49. These results suggest that graph coherence is particularly influential in moving consumers from passive exposure to exploratory product action.

Recognition accuracy also improved under high graph coherence. Participants correctly recognized the endorsed product in 78.4% of high-affinity exposures, 72.6% of category-only exposures, and 68.1% of low-affinity exposures. The effect was stronger when the endorsed product appeared in the first card position, but it remained statistically detectable when the endorsed product appeared in the third or fourth position. This indicates that graph coherence did not merely increase recognition by making the endorsed product more visible. It likely improved encoding by making the product context easier to interpret.

The mixed-effects model confirms these descriptive results. In the purchase-intention model, the virtual-influencer coefficient was -0.24 in the category-only reference condition, indicating lower purchase intention for virtual sources when graph context was not high affinity. The high-affinity graph condition had a positive coefficient of 0.41, while the low-affinity condition had a negative coefficient of -0.28 relative to category-only. The interaction between virtual source and high-affinity condition was positive at 0.19. This means that the virtual-source penalty was smaller when the endorsed product appeared in a coherent graph-based product neighborhood. The interaction between virtual source and low-affinity condition was negative at -0.11, indicating that incoherent panels made virtual-source endorsements somewhat less effective.

The magnitude of the interaction is substantively meaningful. In the category-only condition, the adjusted difference between human and virtual influencers was 0.24 purchase-intention points. In the high-affinity condition, the adjusted difference fell to 0.05 points. In the low-affinity condition, the adjusted difference increased to 0.35 points. Thus, graph coherence nearly eliminated the virtual-source penalty in purchase intention, while graph incoherence amplified it. The same pattern appeared for endorsement fit and skepticism. High-affinity panels reduced skepticism for virtual influencer exposures by 0.22 points more than for human influencer exposures. This supports the interpretation that coherent product neighborhoods provide contextual support for artificial sources.

Trust and authenticity were less fully compensated by graph coherence. Human influencers retained higher trust than virtual influencers in all graph conditions. In the high-affinity condition, the adjusted trust difference was 0.31 points. In the category-only condition, it was 0.49 points. In the low-affinity condition, it was 0.58 points. Graph coherence therefore narrowed the trust gap but did not remove it. Authenticity showed a similar pattern. This distinction matters because purchase intention can improve even when source credibility gaps remain. A coherent product context may help consumers consider the product despite still rating the virtual source as less authentic.

Novelty showed the opposite pattern. Virtual influencers had higher novelty in every graph condition, but high-affinity panels slightly reduced the novelty gap. In the low-affinity condition, virtual influencers were rated 0.84 points higher in novelty than human influencers. In the high-affinity condition, the difference was 0.63 points. This likely occurred because high-affinity panels shifted attention from the source as an unusual object to the product routine as a meaningful context. In other words, graph coherence made the advertisement less dependent on source novelty.

Panel-level click behavior supports the same conclusion. In the high-affinity condition, participants clicked either the endorsed product or a strongly related product in 31.8% of exposures. In the category-only condition, the rate was 24.6%. In the low-affinity condition, the rate was

Table 9. Structural Response Model Coefficients

Predictor	Coefficient	Effect
Trust	0.42	Positive
Endorsement Fit	0.31	Positive
Panel Coherence	0.27	Positive
Novelty	0.13	Positive
Skepticism	-0.24	Negative
Virtual Source (Adjusted)	-0.04	Weak

Table 10. Robustness and Diagnostic Tests

Robustness Test	Main Effect	Interaction	Result
Top 15% Threshold	0.38	Positive	Stable
Top 20% Threshold	0.34	Positive	Stable
No PMI Weight	0.36	0.16	Similar Pattern
No Routine Consistency	0.32	Positive	Reduced Magnitude
Participant Fixed Effects	0.39	0.17	Consistent
Middle Position Only	0.35	0.15	Robust

18.9%. For virtual influencer exposures specifically, the high-affinity condition increased click rate from 22.1% in category-only to 30.7%. For human influencer exposures, the increase was from 27.1% to 32.9%. The absolute click rate remained slightly higher for human sources, but the graph-coherence lift was larger for virtual sources.

The continuous graph-coherence model provides a more refined view. A one-standard-deviation increase in graph coherence was associated with a 0.29-point increase in purchase intention, a 0.33-point increase in product consideration, a 0.37-point increase in perceived panel coherence, and a 0.21-point increase in endorsement fit. The interaction between virtual source and graph coherence was positive for purchase intention, product consideration, and endorsement fit. Category match had a smaller positive coefficient after graph coherence was included. Routine-domain match remained positive, but its coefficient was also smaller than graph coherence. This suggests that the graph-affinity measure captured more than category and routine labels.

The graph-regularized model improved out-of-sample prediction of purchase intention. Ordinary ridge regression using source, product, and participant covariates achieved a test root mean squared error of 1.142. Adding product fixed effects reduced the error to 1.109. The graph-regularized model reduced the error further to 1.073. The improvement is moderate, but it indicates that product-neighborhood smoothing added predictive information. The selected Laplacian penalty was not zero, meaning that the cross-validation process favored some smoothing across connected products. Products in the same graph neighborhood tended to have related response patterns even after controlling for experimental condition.

The high-intention classification models show that source and graph variables jointly improve prediction. The source-only logistic model achieved ROC-AUC of 0.714 and PR-AUC of 0.402. The product-only model achieved ROC-AUC of 0.681 and PR-AUC of 0.371. The graph-only model achieved ROC-AUC of 0.746 and PR-AUC of 0.438. Combining source and graph variables increased ROC-AUC to 0.803 and PR-AUC to 0.521. The full model, which added behavioral variables and perception measures, achieved ROC-AUC of 0.862 and PR-AUC of 0.614 using gradient boosting. Logistic regression on the full feature set achieved ROC-AUC of 0.831 and PR-AUC of 0.571. Random forest achieved ROC-AUC of 0.844 and PR-AUC of 0.589.

Calibration diagnostics favored penalized logistic regression, while discrimination favored gradient boosting. The full logistic model had a Brier score of 0.158 and expected calibration error of 0.036. The boosted model had a Brier score of 0.151 and expected calibration error of 0.052 before calibration. Isotonic calibration reduced the boosted model's expected calibration error to 0.031 and improved the Brier score to 0.147. Therefore, the calibrated boosted model offered the best overall prediction performance. This result is useful for practical campaign diagnostics because calibrated probabilities can be used to compare expected high-intention rates across creative and panel conditions.

Permutation importance in the boosted model ranked perceived panel coherence, trust, graph coherence, product-card expansion, endorsement fit, novelty, source type, skepticism, and panel dwell time as the strongest predictors. Product category, category match, and source identity were less important after graph coherence and perceptions were included. The importance ranking suggests that the con-

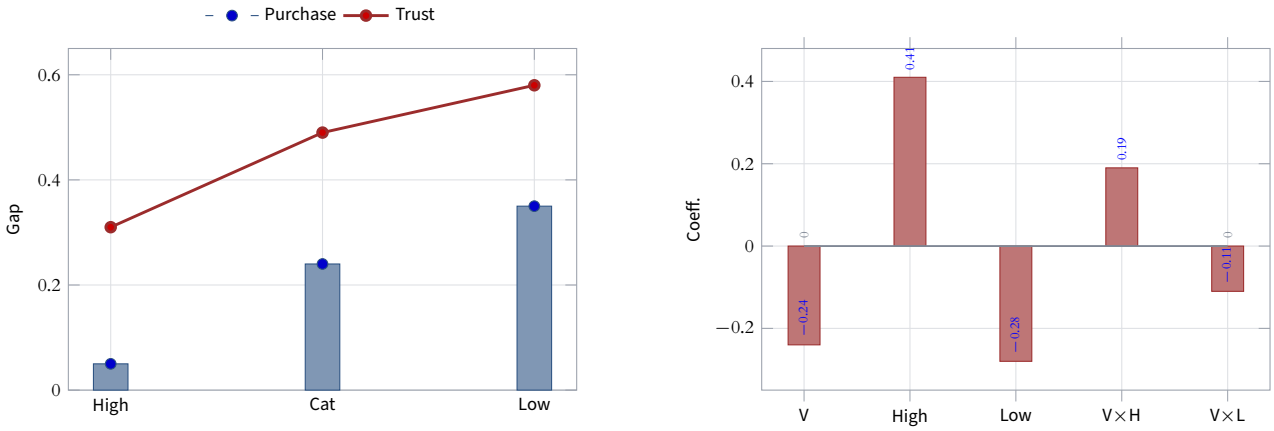


Figure 3. High-affinity neighborhoods nearly eliminated the virtual-source purchase-intention penalty while narrowing, but not removing, the trust gap. The mixed-effects coefficients show that high graph affinity raised purchase intention, low affinity reduced it, and the virtual-source interaction was positive only in coherent product neighborhoods.

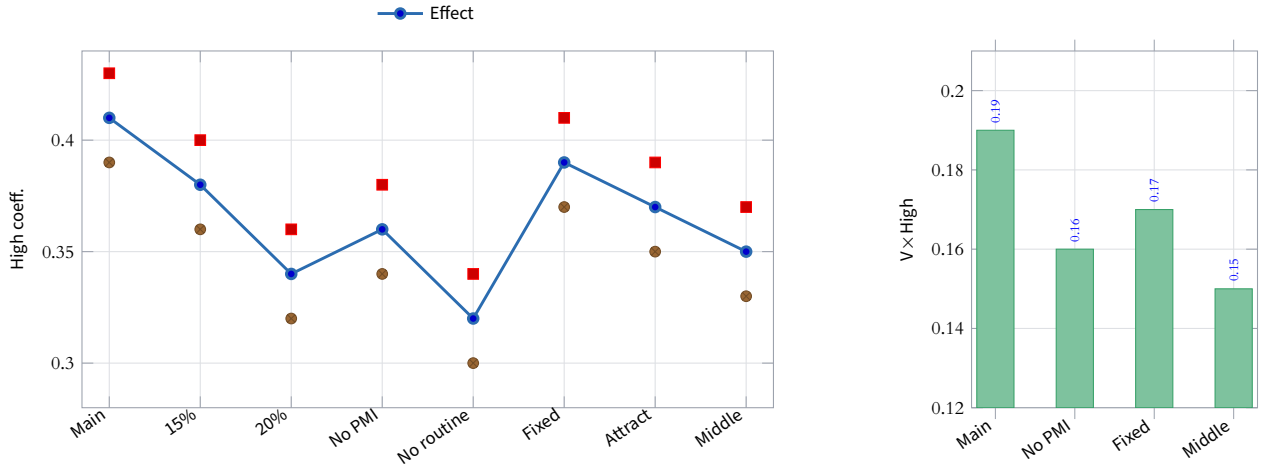


Figure 4. The high-affinity effect remained positive across alternative graph thresholds, graph-weight exclusions, participant fixed effects, attractiveness controls, and middle-position restrictions. The virtual-source interaction also stayed positive in the available robustness specifications, indicating that the penalty reduction was not tied to a single modeling choice.

sumer's interpretation of the product neighborhood is central. The graph variable matters directly, but its perceived counterpart matters even more. This is reasonable because the graph structure affects behavior only insofar as it creates a meaningful visible context for the consumer.

The structural response model shows that source type operates largely through perceptions. In a model containing only source type, graph condition, and controls, the virtual-source coefficient for purchase intention was -0.21. After adding trust, authenticity, novelty, skepticism, endorsement fit, and panel coherence, the coefficient became -0.04 and was not statistically distinguishable from zero. Trust had a coefficient of 0.42, endorsement fit had a coefficient of 0.31, panel coherence had a coefficient of 0.27, skepticism had a coefficient of -0.24, and novelty had a coefficient of 0.13. Authenticity remained positive but smaller after trust was included. This pattern indicates that virtual-source effects are mostly explained by lower trust, lower authenticity, higher novelty, and changes in perceived coherence.

The model for click intention differed from the model for purchase intention. Novelty had a larger coefficient for click intention than for purchase intention. Product-card expansion and panel dwell time were also stronger predictors of click intention. Trust remained positive but was less dominant. This distinction suggests that virtual influencers may be more effective at generating exploration than commitment. When graph coherence is high, that exploration can be directed toward relevant products rather than dissipating as mere curiosity about the artificial source.

Explicit disclosure had mixed but interpretable effects. It slightly reduced sponsored-post dwell time by 0.08 seconds, suggesting that clear commercial labeling may reduce some initial inspection. However, it increased trust by 0.17 points and reduced skepticism by 0.19 points. For virtual influencers, explicit disclosure increased trust by 0.24 points and reduced skepticism by 0.27 points. Disclosure also strength-

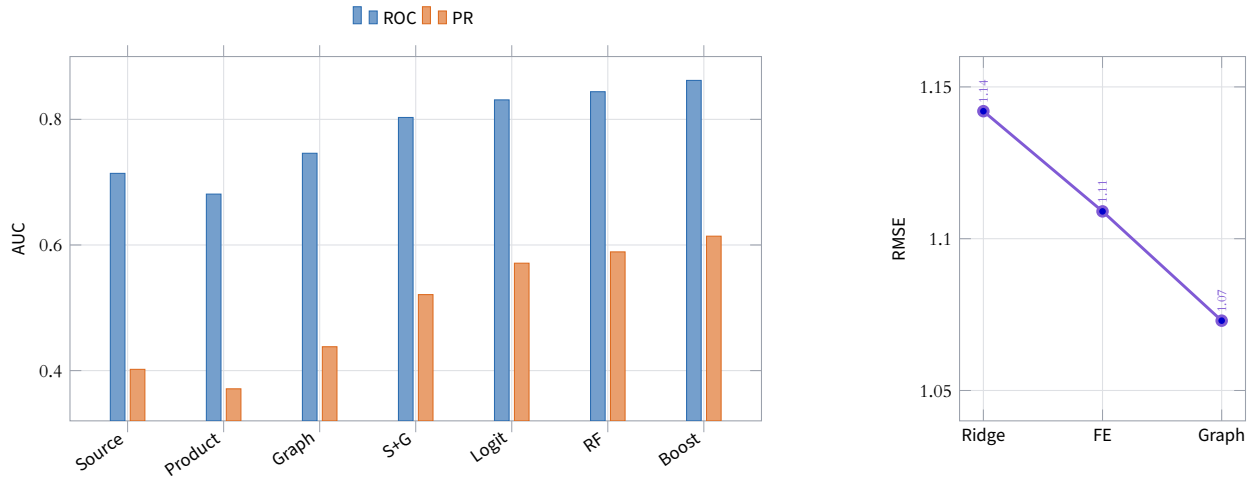


Figure 5. High-intention classification improved when source and graph variables were combined, and the full gradient-boosting model achieved the strongest discrimination. In the regression diagnostic, graph regularization reduced test RMSE beyond ordinary ridge regression and product fixed effects, showing that product-neighborhood smoothing added predictive information.

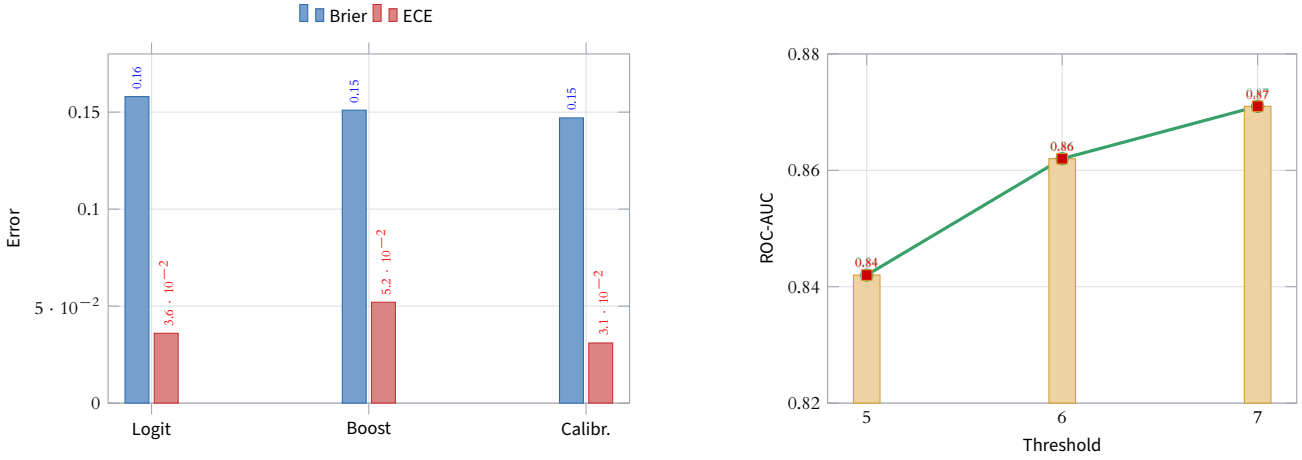


Figure 6. Calibration diagnostics favored penalized logistic regression before calibration, while discrimination favored gradient boosting. Isotonic calibration reduced the boosted model's expected calibration error and improved the Brier score, and alternative high-intention thresholds preserved strong ROC-AUC performance.

ened the positive effect of graph coherence on endorsement fit. This suggests that transparency and contextual coherence work together. When the source is clearly identified and the product neighborhood makes sense, the endorsement appears less arbitrary.

The results therefore support the paper's main claim. Influencer advertising in product discovery environments is shaped by both source credibility and product-neighborhood coherence. Human influencers retain advantages in trust and authenticity. Virtual influencers retain advantages in novelty and initial attention. High graph-affinity panels improve recognition, consideration, click intention, and purchase intention. Most importantly, high graph-affinity panels reduce the purchase-intention penalty associated with virtual influencers. The product graph does not erase source differences, but it changes how those differences translate into response.

6. Robustness and Diagnostic Tests

Several robustness tests were conducted to assess whether the results depend on specific modeling choices. The first test replaced the high-affinity graph condition with alternative graph thresholds. The main analysis used the top decile of edge weights for high-affinity panel construction. Alternative versions used the top 15% and top 20% of edges. The estimated high-affinity effect on purchase intention was 0.41 in the main model, 0.38 with the top 15% threshold, and 0.34 with the top 20% threshold. The virtual-source by high-affinity interaction remained positive in all versions. The declining coefficient is expected because loosening the threshold admits weaker product relationships.

The second test excluded pointwise mutual information from the edge-weight formula. PMI can overstate rare associations even when minimum-count filters are used. The alternative edge weight combined support, confidence, lift,

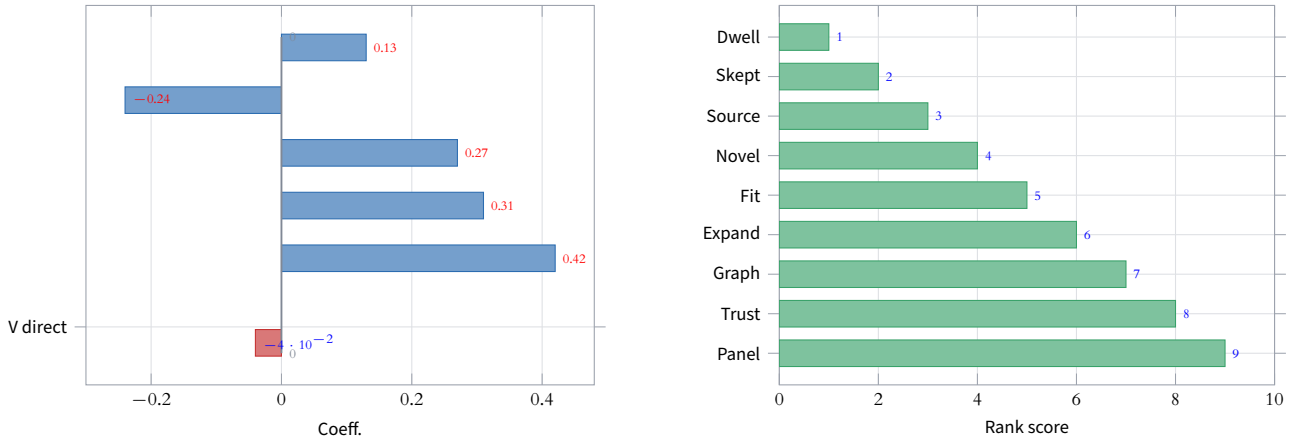


Figure 7. The structural response model shows that source effects were largely mediated by trust, endorsement fit, panel coherence, skepticism, and novelty. The permutation-order diagnostic places perceived panel coherence, trust, graph coherence, product-card expansion, and endorsement fit among the strongest predictive drivers.

and routine consistency. The high-affinity coefficient declined from 0.41 to 0.36, and the virtual-source interaction declined from 0.19 to 0.16. The results remained substantively similar. This suggests that the graph-coherence effect is not driven solely by rare high-PMI pairs.

The third test removed routine-domain consistency from the graph weight. This was a stricter test because routine labels were part of the stimulus-design logic. Without the routine component, the high-affinity effect remained positive at 0.32, and perceived panel coherence remained substantially higher than in the low-affinity condition. However, the difference between high-affinity and category-only conditions became smaller. This indicates that behavioral co-occurrence alone has value, but routine-consistent cross-category relations contribute to the strongest consumer interpretation.

The fourth test used ordered logit models for seven-point outcomes. The ordered models produced the same sign pattern as the linear mixed-effects models. High-affinity condition increased the odds of higher purchase-intention categories, virtual influencer source decreased the odds in the category-only reference condition, and the virtual-source by high-affinity interaction increased the odds. Because the linear models are easier to interpret in scale points and are stable under cluster-robust estimation, they remain the main reported models.

The fifth test adjusted for participant response style. Some participants used high scale values across most items, while others used lower values. To address this, purchase intention and perception measures were standardized within participant before estimating the structural model. Trust, endorsement fit, panel coherence, and skepticism remained the strongest predictors. The direct virtual-source coefficient remained close to zero after perceptions were included. This suggests that the structural interpretation is not an artifact of general response positivity.

The sixth test used participant fixed effects instead of participant random effects. Because each participant saw multiple exposures, fixed effects provide a stricter within-participant comparison. The high-affinity coefficient remained positive at 0.39, and the virtual-source by high-affinity interaction was 0.17. The virtual-source main effect was slightly smaller in magnitude than in the random-effects model. This indicates that between-participant differences are not driving the core findings.

The seventh test excluded participants who correctly guessed that the study compared human and virtual influencers. At the end of the session, participants wrote a short description of what they thought the study examined. Approximately 16.8% mentioned human versus AI or virtual influencers. Excluding these participants reduced the virtual-source novelty advantage and trust penalty slightly, but the high-affinity interaction remained. This suggests that demand awareness may sharpen source-type judgments, but it does not explain the graph-coherence effect.

The eighth test examined whether the high-affinity effect was merely a consequence of showing more attractive products. The product pool was balanced so that each product appeared across graph conditions, but the panel products differed by condition. To address this, models included average product attractiveness scores from an independent pretest. The high-affinity coefficient declined only slightly, from 0.41 to 0.37. Pretest attractiveness was positive, but it did not eliminate the graph effect. This supports the interpretation that relational coherence matters beyond product appeal.

The ninth test controlled for visual similarity among products. High-affinity panels might appear more coherent because products look similar, not because they are graph related. A computer-vision embedding was used to compute average visual similarity between the endorsed product and panel products. Including this variable reduced the category-match coefficient but did not eliminate graph coherence. The graph-coherence coefficient remained 0.26 in the continuous model. This indicates that behavioral affinity is not reducible to visual similarity.

The tenth test evaluated potential position effects. The endorsed product appeared in either first position or middle position. When the model was estimated only on middle-position exposures, the high-affinity coefficient remained positive at 0.35, and the virtual-source by high-affinity interaction remained 0.15. This shows that high-affinity effects are not solely caused by first-position visibility. Panel context still mattered when the endorsed product was not placed first.

The eleventh test examined source-identity heterogeneity. There were twelve human influencer identities and twelve virtual influencer identities. Random slopes for source type by identity showed variation. Some virtual identities performed nearly as well as human identities in trust, especially when their visual style was less stylized and disclosure was explicit. Some human identities performed poorly when product fit was weak. Nevertheless, the average source-type pattern remained. This indicates that source type matters, but identity design and product fit also matter.

The twelfth test examined product heterogeneity. The strongest graph-coherence effects appeared for travel pouch, compact backpack, coffee grinder, and notebook set. These products have obvious routine complements. The weakest effects appeared for skincare serum and wireless earbuds, where evaluation may depend more on technical claims, reviews, or personal preference. This variation suggests that graph-affinity context is most useful when the product belongs to a multi-item routine. Future work should test more product categories to estimate where graph context is most persuasive.

The thirteenth diagnostic checked predictive leakage. The full prediction model includes perception variables that are measured after exposure, so it is useful for diagnosis but not for real-time pre-exposure prediction. A restricted pre-response model using only source type, product variables, graph variables, disclosure, panel order, and behavioral logs available before the survey achieved ROC-AUC of 0.804. This is lower than the full model's 0.862 but still meaningfully above the source-only and graph-only baselines. The restricted model is more relevant for live campaign optimization, while the full model is more relevant for explaining response.

The fourteenth diagnostic examined calibration across source type. Some models were better calibrated for human influencer exposures than for virtual influencer exposures. The uncalibrated boosted model overpredicted high purchase intention for virtual influencer exposures in the low-affinity condition. After isotonic calibration within validation folds, this bias declined [31]. Calibration by source type is important because a model that systematically overestimates virtual-source performance could mislead campaign allocation.

The fifteenth diagnostic used alternative high-intention thresholds. The main high-intention label used purchase intention of at least 6. When the threshold was changed to at least 5, the positive class increased and the boosted model achieved ROC-AUC of 0.842. When the threshold was restricted to 7 only, the positive class became smaller and the model achieved ROC-AUC of 0.871 but lower PR-AUC because positives were rarer. The feature-importance ranking remained similar, with panel coherence, trust, graph coherence, and endorsement fit near the top.

These robustness checks support the main findings. The graph-coherence effect remains under alternative graph weights, model forms, participant controls, visual-similarity controls, position restrictions, and threshold changes. The virtual-source penalty is consistently reduced in high-affinity product neighborhoods. The effect is not large enough to claim that graph context fully substitutes for source credibility, but it is consistent enough to suggest that product-neighborhood design is a meaningful part of influencer advertising effectiveness.

7. Discussion

The findings show that influencer advertising in e-commerce product discovery should be understood as a joint source-context process. Human influencers and virtual influencers differ in credibility profiles, but those differences do not operate in isolation. The product neighborhood surrounding the endorsement changes how the consumer interprets the promoted item. When the product panel is graph coherent, the endorsement appears more meaningful, recognition improves, and purchase intention increases. When the panel is incoherent, the endorsement appears more isolated and skepticism increases, especially for virtual influencers.

The most important empirical result is the interaction between virtual source type and high graph affinity. Virtual influencers were rated lower in trust and authenticity than human influencers across conditions, but their purchase-intention disadvantage was much smaller when the surrounding product panel was graph coherent. This suggests that product-neighborhood evidence can partly compensate for weaker source credibility. The compensation is not complete. Human influencers still retain trust advantages. However, a coherent graph context gives the consumer a reason to understand the product recommendation as part of a routine rather than as a detached artificial promotion.

This result has implications for how virtual influencers should be evaluated. A virtual influencer may be a poor fit for campaigns that depend mainly on personal testimony, embodied experience, or claims requiring human use. But the same virtual influencer may perform adequately when the campaign is embedded in a coherent product-discovery environment that supplies contextual relevance. In such cases, the virtual persona attracts attention and the product graph supplies interpretability. The campaign works less like a personal recommendation and more like a stylized entry point into a product routine.

The findings also qualify simple claims about novelty. Virtual influencers produced higher novelty and slightly stronger initial engagement, but novelty was more strongly associated with click intention than with purchase intention.

This means novelty can generate exploration without guaranteeing commitment. High-affinity product panels help convert some of that exploration into product consideration by directing attention toward relevant complements. Low-affinity panels leave novelty unsupported, which can make the artificial source appear more like an attention device than a credible recommender.

For human influencers, graph coherence also improved response, but the mechanism was somewhat different. Human influencers already had higher trust and authenticity, so graph coherence mainly strengthened product fit and panel usefulness. The endorsement did not require the graph to appear credible, but it benefited from the graph when the product neighborhood helped participants imagine a usage context. Human-source campaigns may therefore use graph-coherent panels to improve basket expansion and cross-sell behavior, while virtual-source campaigns may need graph coherence to reduce perceived arbitrariness.

The study also contributes to product-discovery measurement. Category similarity is not enough to describe product context. The category-only condition performed better than low affinity but worse than high affinity. This means that products from the same category can be understandable but not necessarily persuasive as a recommendation neighborhood. Graph-affinity panels can include cross-category complements that better represent actual routines. A travel pouch may be more meaningfully paired with packing cubes, passport holders, and compact chargers than with only other pouches. A coffee grinder may be more meaningfully paired with filters, beans, and a notebook than with only other grinders. These relations are difficult to capture with category labels alone.

The graph-regularized results support this point statistically. Product-neighborhood smoothing improved prediction beyond ordinary ridge regression and product fixed effects. This does not mean that the graph model is the only or best way to represent product context, but it indicates that graph proximity captures response-relevant structure. In e-commerce advertising, product relationships can be persuasive because they help consumers form a mental model of use. A graph-coherent panel tells a small story about how products belong together.

The predictive results show that practical campaign diagnostics benefit from combining source, graph, perception, and behavior variables. The full boosted model achieved the strongest high-intention classification, but the source-plus-graph model already outperformed source-only and product-only baselines. This suggests that product-neighborhood design should be part of campaign testing, not an afterthought handled separately by a recommender system. If an influencer campaign sends consumers into an incoherent product panel, some persuasive value may be lost. If the panel reinforces the endorsed product's routine, the same creative may perform better.

There are also ethical and transparency implications. Explicit disclosure reduced skepticism and improved trust, especially for virtual influencers. This suggests that clear identification of artificial sources does not necessarily weaken campaign response. It can improve the interpretive environment by reducing ambiguity. When disclosure is paired with graph coherence, the virtual endorsement appears less deceptive and more contextually grounded. This is important because artificial personas can blur the boundary between entertainment, recommendation, and brand-controlled persuasion. Transparent labeling should be treated as part of credible design rather than as a legal formality alone.

The study's results should not be overstated. The experiment used controlled stimuli, simulated product graphs, and immediate response outcomes. Real platforms contain richer histories, comments, ratings, dynamic prices, social proof, and personalization. A live recommender system may optimize for engagement in ways that differ from the graph construction used here. Nevertheless, the controlled design isolates a relationship that would be difficult to observe cleanly in platform data: the interaction between endorser type and product-neighborhood coherence. The evidence suggests that this interaction is real enough to justify further field testing.

For future advertising research, the study points toward a broader unit of analysis. The sponsored post alone is too narrow. The full persuasive object includes the source, the product, the disclosure, the adjacent recommendations, the product graph, and the behavioral path that follows exposure. Measuring only source perception misses the product-network contribution. Measuring only recommendation clicks misses the persuasive role of the source. A joint model is needed because modern e-commerce discovery is jointly social, algorithmic, and relational.

8. Conclusion

This paper examined how influencer source type and graph-based product-affinity context jointly shape consumer response in mobile e-commerce advertising. The study used a controlled experiment with 1,536 adult participants, 6,144 sponsored exposures, and 49,152 product-panel observations. Human influencers produced higher trust, authenticity, and parasocial response, while AI-generated virtual influencers produced higher novelty and slightly stronger initial feed interruption. These source differences were meaningful, but they did not fully determine purchase intention.

The product-neighborhood context substantially changed consumer response. High-affinity graph panels improved perceived panel coherence, product recognition, product consideration, click intention, and purchase intention relative to category-only and low-affinity panels. This indicates that product context is not merely a visual background for the endorsement. It helps consumers interpret whether the promoted product belongs to a meaningful routine. Category similarity helped, but graph-based affinity helped more, especially when cross-category products formed a coherent usage context.

The interaction between source type and graph co-

herence was the central finding. Virtual influencers had a purchase-intention disadvantage in category-only and low-affinity contexts, but that disadvantage became very small in high-affinity contexts. Graph coherence did not eliminate the trust and authenticity gap between virtual and human influencers, but it reduced the extent to which that gap translated into lower purchase intention. This suggests that coherent product-neighborhood evidence can partly support artificial sources by making the endorsed product appear less arbitrary.

The predictive models reinforced the experimental results. Source-only and product-only models were weaker than models that included graph variables, and the strongest model combined source perceptions, graph coherence, behavioral engagement, and product context. A calibrated gradient boosting model provided useful discrimination and probability estimates for high-intention exposures. The feature-importance results showed that perceived panel coherence, trust, graph coherence, endorsement fit, and skepticism were among the strongest predictors. These variables describe a joint persuasion environment rather than a single creative feature.

The study has practical implications for campaign design. Human influencer campaigns may benefit from graph-coherent product panels because they can expand consideration and support cross-selling. Virtual influencer campaigns may need graph coherence even more because it helps compensate for weaker perceived authenticity and trust. Clear disclosure also appears useful, particularly for artificial sources, because it reduces ambiguity and skepticism. The evidence therefore supports a design principle based on alignment: the source, product, disclosure, and recommendation neighborhood should tell a coherent story.

The findings should be interpreted within the limits of the controlled design. The experiment measured immediate response, not long-term sales or repeated exposure. The product graph was constructed for research control, not taken from a live commercial platform. The virtual personas were newly introduced rather than established public figures. Future work should test similar models in field settings, with live recommendation systems, repeated exposures, and behavioral purchase outcomes. Even with these limitations, the results show that influencer advertising effectiveness in digital commerce depends not only on who endorses a product, but also on whether the product's surrounding graph makes the endorsement make sense.

References

[1] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, pp. 207–216, ACM, 1993.

[2] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," *Proceedings of the 10th International Conference on World Wide Web*, pp. 285–295, 2001.

[3] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80, 2003.

[4] C. I. Hovland and W. Weiss, "The influence of source credibility on communication effectiveness," *Public Opinion Quarterly*, vol. 15, no. 4, pp. 635–650, 1951.

[5] R. Ohanian, "Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness," *Journal of Advertising*, vol. 19, no. 3, pp. 39–52, 1990.

[6] K. Freberg, K. Graham, K. McGaughey, and L. A. Freberg, "Who are the social media influencers? a study of public perceptions of personality," *Public Relations Review*, vol. 37, no. 1, pp. 90–92, 2011.

[7] M. De Veirman, V. Cauberghe, and L. Hudders, "Marketing through instagram influencers: The impact of number of followers and product divergence on brand attitude," *International Journal of Advertising*, vol. 36, no. 5, pp. 798–828, 2017.

[8] C. Lou and S. Yuan, "Influencer marketing: How message value and credibility affect consumer trust of branded content on social media," *Journal of Interactive Advertising*, vol. 19, no. 1, pp. 58–73, 2019.

[9] B. F. Török, "A validated synthetic testbed for graph-based product affinity research in e-commerce," *Transactions on Artificial Intelligence, Machine Learning, and Cognitive Systems*, vol. 9, no. 8, pp. 1–25, 2024.

[10] B. F. Török, "Comparative effects of virtual and human influencer endorsements on purchase intent among digital consumer audiences in lifestyle marketing," *Northern Reviews on Algorithmic Research, Theoretical Computation, and Complexity*, vol. 10, no. 11, pp. 1–20, 2025.

[11] E. Moustakas, N. Lamba, D. Mahmoud, and C. Ranganathan, "Blurring lines between fiction and reality: Perspectives of experts on marketing effectiveness of virtual influencers," in *2020 International Conference on Cyber Security and Protection of Digital Services*, pp. 1–6, IEEE, 2020.

[12] S. Sands, C. L. Campbell, K. Plangger, and C. Ferraro, "Unreal influence: Leveraging ai in influencer marketing," *European Journal of Marketing*, vol. 56, no. 6, pp. 1721–1747, 2022.

[13] V. L. Thomas and K. Fowler, "Close encounters of the ai kind: Use of ai influencers as brand endorsers," *Journal of Advertising*, vol. 50, no. 1, pp. 11–25, 2021.

[14] M. Friestad and P. Wright, "The persuasion knowledge model: How people cope with persuasion attempts," *Journal of Consumer Research*, vol. 21, no. 1, pp. 1–31, 1994.

[15] S. C. Boerman, L. M. Willemsen, and E. P. Van Der Aa, ""this post is sponsored": Effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of facebook," *Journal of Interactive Marketing*, vol. 38, pp. 82–92, 2017.

[16] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pp. 452–461, AUAI Press, 2009.

[17] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, eds., *Recommender Systems Handbook*. Boston, MA: Springer, 2011.

[18] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations*, 2017.

[19] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[20] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *International Conference on Learning Representations*, 2018.

[21] X. Wang, X. He, M. Wang, F. Feng, and T.-S. Chua, "Neural graph collaborative filtering," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 165–174, 2019.

[22] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "Lightgcn: Simplifying and powering graph convolution network for recommendation," in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 639–648, 2020.

[23] J. R. Bettman, M. F. Luce, and J. W. Payne, "Constructive consumer choice processes," *Journal of Consumer Research*, vol. 25, no. 3, pp. 187–217, 1998.

[24] J. Alba, J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, and S. Wood, "Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces," *Journal of Marketing*, vol. 61, no. 3, pp. 38–53, 1997.

[25] G. Häubl and V. Trifts, "Consumer decision making in online shopping environments: The effects of interactive decision aids," *Marketing Science*, vol. 19, no. 1, pp. 4–21, 2000.

[26] P. Chandon, J. W. Hutchinson, E. T. Bradlow, and S. H. Young, "Does in-store marketing work? effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase," *Journal of Marketing*, vol. 73, no. 6, pp. 1–17, 2009.

[27] L. Breiman, "Statistical modeling: The two cultures," *Statistical Science*, vol. 16, no. 3, pp. 199–231, 2001.

[28] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* New York: Springer, 2 ed., 2009.

[29] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[30] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[31] N. V. Patel, "Selection bias in two-sided e-commerce marketplaces: A framework for propensity score matching implementation," *Journal of Business Intelligence Systems and Computational Social Science Applications*, vol. 11, no. 5, pp. 1–20, 2021.