

Article

Short-Form Video Marketing: Relationships with Consumer Perceptions and Purchase Intention

Galina Ilieva ^{1,*}, Tania Yankova ¹, Margarita Ruseva ¹, Delian Angelov ², Stanislava Klisarova-Belcheva ¹, Marin Bratkov ³, Penyo Georgiev ¹ and Angel Dimitrov ¹

¹ Department of Management and Quantitative Methods in Economics, University of Plovdiv Paisii Hilendarski, 4000 Plovdiv, Bulgaria; tsvl@uni-plovdiv.bg (T.Y.); m_ruseva@uni-plovdiv.bg (M.R.); stanislava.belcheva@uni-plovdiv.bg (S.K.-B.); penyo@uni-plovdiv.bg (P.G.); angel.dimitrov@uni-plovdiv.bg (A.D.)

² Department of Economic Sciences, University of Plovdiv Paisii Hilendarski, 4000 Plovdiv, Bulgaria; da@uni-plovdiv.bg

³ Department of Marketing and International Economic Relations, University of Plovdiv Paisii Hilendarski, 4000 Plovdiv, Bulgaria; m_bratkov@uni-plovdiv.bg

* Correspondence: galili@uni-plovdiv.bg

Highlights

Please indicate how your work links to systems science via your contributions to systems practice, theory, and or methodology.

- The study applies a systems-oriented BI methodology to data from 409 Bulgarian consumers by integrating descriptive profiling, formative PLS path modeling, group-difference analysis, clustering, sentiment analysis, and machine-learning prediction.
- It represents consumer responses to short-form video marketing as an interconnected system in which five content- and creator-related perceptions are associated with Creating Shared Value (CSV), formed by eight distinct product/economic, community/social, and relational contributions, and subsequently with purchase intention.

What are the main findings and/or the implications of the main findings?

- Clarity, willingness to use, similarity, likability, and empathy are positively associated with CSV, which is positively associated with purchase intention.
- Age and followed-influencer count distinguish mean CSV and purchase intention, while social-media use frequency and daily SFV viewing time distinguish purchase intention only; clustering, sentiment analysis, and machine learning provide complementary BI outputs for segmentation, text analysis, and prediction.

Abstract

Short-form video (SFV) marketing generates structured consumer ratings and unstructured feedback that can be analyzed for e-commerce decision support. This study examines associations between five content- and creator-related perceptions, Creating Shared Value (CSV), and purchase intention among 409 respondents in Bulgaria. From a business intelligence (BI) perspective, this study combines descriptive profiling, clustering, sentiment analysis, formative partial least squares (PLS) path modeling, group-difference testing, and machine-learning prediction to produce descriptive profiles, explanatory associations, consumer segments, text-analytic findings, and predictive results. CSV is specified as a formative composite of eight noninterchangeable product/economic, community/social, and relational indicators associated with short-form video creators. Clarity ($\beta = 0.260$), willingness to use ($\beta = 0.226$), similarity ($\beta = 0.205$), likability ($\beta = 0.177$), and empathy ($\beta = 0.130$) are positively associated with CSV; CSV is positively associated with purchase



Academic Editors: Jong-min Kim and Rob Kim Marjerison

Received: 5 July 2026

Revised: 29 July 2026

Accepted: 30 July 2026

Published: 1 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

intention ($\beta = 0.567$). The model explains 69.1% of CSV variance and 32.2% of purchase-intention variance. Six formative weights are significant, all absolute loadings exceed 0.72, and indicator VIFs remain below 5. The group-difference hypothesis is partially supported by 18 omnibus ANOVAs with Holm correction for multiple testing. Significant differences in both CSV and purchase-intention means are found across age groups and categories based on the number of influencers followed. Significant differences in purchase-intention means are also found across social-media use-frequency and daily SFV-viewing-time categories. Clustering, sentiment-analysis, and machine-learning results are interpreted as complementary BI outputs rather than evidence of causal effects or validated campaign effectiveness.

Keywords: short-form video marketing; e-commerce; business intelligence (BI); consumer analytics; Creating Shared Value (CSV); formative measurement; purchase intention; partial least squares structural equation modeling (PLS-SEM); clustering; sentiment analysis; machine learning

1. Introduction

Short-form video (SFV) marketing has become a prominent component of e-commerce communication. Mobile-first feeds combine compressed audiovisual content, creator mediation, recommendation mechanisms, and immediate interactions, including likes, comments, saves, shares, and product-page visits. These features enable short videos to function simultaneously as advertising, product information, social influence, and a source of consumer-response data [1–3].

Research associates the persuasive value of SFV with message usefulness, clarity, entertainment, trust, product visualization, and creator-related perceptions [4,5]. Unlike conventional video advertising, SFV is encountered in rapid, algorithmically curated sequences in which users can move from exposure to information search, engagement, or purchase-related action with minimal friction. This combination makes SFV relevant to both consumer-behavior research and BI-supported e-commerce analysis.

Prior research links SFV characteristics to trust, advertising value, attitude, shared value, and purchase intention [4–7]. However, this evidence is fragmented across separate explanatory models, concentrated on particular platforms and national contexts, and rarely combines structured ratings, text analysis, segmentation, and predictive analytics within a single study. Research on European e-commerce users, especially those in Central and Eastern Europe, remains comparatively limited.

The present study addresses this gap using survey data from Bulgarian online consumers. Its primary model follows the stimulus–organism–response (SOR) logic: clarity, likability, empathy, similarity, and willingness to use serve as antecedent evaluations; Creating Shared Value (CSV) represents the focal internal value appraisal; and self-reported purchase intention represents the response [8,9]. Technology-acceptance and related perspectives contextualize particular antecedents but are not treated as competing primary models [10,11].

This study tests whether clarity, likability, empathy, similarity, and willingness to use are positively associated with CSV and whether CSV is positively associated with purchase intention. It also examines whether the prespecified demographic and usage groups differ in their mean CSV-indicator and/or purchase-intention scores.

Creating Shared Value concerns the joint creation of economic and social value [7]. In this study, CSV is measured as a formative composite of consumers' perceptions of the

product/economic, community/social, and relational contributions associated with short-form video creators. Respondents cannot independently verify objective sales, employment, or community outcomes. Accordingly, CSV scores are interpreted as perception-based assessments rather than direct evidence of objectively realized value.

Conceptually, this study extends the SOR interpretation of SFV marketing by representing the organism component as a formative value appraisal composed of product/economic, community/social, and relational contributions. CSV is particularly suitable for the present research question because these noninterchangeable contributions jointly define the value that consumers attribute to short-video creators, whereas trust and attitude primarily capture confidence in, or general favorability toward, a creator or marketing stimulus. The framework also refines the role of technology-acceptance reasoning: willingness to use is not treated as the final behavioral outcome but as an antecedent evaluation reflecting consumers' readiness to employ SFV content as an informational or decision-support resource. It is therefore positioned within the SOR sequence as an evaluation associated with the formation of CSV, while purchase intention represents the subsequent response. Accordingly, the framework does more than combine SOR and TAM constructs; it distinguishes technology-related readiness, multidimensional value appraisal, and purchase intention as conceptually separate stages. This study does not claim that CSV is universally superior to trust, attitude, or perceived value; rather, it offers a broader value-oriented explanation for the context investigated.

The analysis addresses four research questions:

RQ1: Which content- and creator-related perceptions are most strongly associated with CSV?

RQ2: How strongly is CSV associated with self-reported purchase intention?

RQ3: What complementary segmentation, text-analytic, and predictive outputs are produced by the BI analyses?

RQ4: Across which prespecified demographic and usage classifications do mean CSV-indicator and/or purchase-intention scores differ?

The BI-oriented multimethod approach combines descriptive profiling, hierarchical and k-means clustering, sentiment analysis, a partial least squares path model in which CSV is specified formatively, group-difference analysis, and machine-learning prediction. The primary PLS model evaluates the hypothesized structural associations. The complementary procedures provide consumer segmentation, text-analytic interpretation, group-difference assessment, and internal predictive evaluation. Their outputs are not treated as components of a single confirmatory model [12,13].

The BI contribution lies in integrating structured survey indicators and open-ended feedback to produce descriptive profiles, explanatory associations, consumer segments, text-analytic findings, and predictive results relevant to e-commerce. The study does not present an implementation-ready organizational workflow. Operational deployment would require external validation using campaign, clickstream, and sales data.

This study makes three bounded contributions. First, it provides evidence from Bulgaria concerning associations between SFV perceptions, CSV, and purchase intention. Second, it respecifies CSV as a formative composite and evaluates its formative weights, loadings, collinearity, and robustness. Third, it demonstrates how complementary BI methods can connect explanatory association, segmentation, text analysis, group comparison, and prediction without conflating association with causality or internal prediction with campaign effectiveness.

From a systems perspective, the analysis connects several measured components of the consumer-response process: content and creator evaluations, a formative CSV composite, purchase intention, consumer heterogeneity, and BI outputs. Platform algorithms, actual

exposure, creator behavior, and realized sales remain contextual components rather than variables directly tested through the survey.

The remainder of the paper presents the theoretical background and hypotheses (Section 2), methodology (Section 3), results (Section 4), discussion (Section 5), and conclusions (Section 6).

2. State-of-the-Art Review in Short Video Marketing

SFV marketing is a mobile, feed-based form of digital persuasion embedded in social-commerce and e-commerce systems. Its effectiveness depends on content value, creator cues, platform affordances, and consumer characteristics rather than video length alone [1–6].

Three groups of factors recur in the literature: content characteristics such as usefulness, clarity, entertainment, and product visualization; creator and relational characteristics such as credibility, empathy, likability, and similarity; and platform or usage characteristics such as recommendation, interactivity, social presence, and viewing intensity [14–18].

In BI analysis, these factors produce different marketing objects: behavioral and perceptual profiles for descriptive analysis, respondent groups for segmentation, open comments for text analytics, construct relationships for explanation, and response indicators for prediction. The present paper integrates these outputs while keeping their evidential roles separate.

2.1. SFV Marketing as an E-Commerce and BI Context

SFV combines rapid attention capture, informational and entertainment value, creator-mediated social cues, interaction tools, and embedded commerce. These characteristics shorten the path between exposure, information search, evaluation, engagement, and purchase-related action, while generating data that e-commerce firms can analyze at the content, audience, and campaign levels [1–6,14–18].

The available data support several BI functions: descriptive profiling of audiences and usage, diagnostic explanation of construct relationships, segmentation of response patterns, sentiment analysis of qualitative feedback, and predictive modeling of value-related responses. These analytical functions motivate the multi-method design used in this study; they do not by themselves prove operational campaign effects.

2.2. Primary Theoretical Framework and Contextual Perspectives

The SOR framework serves as the organizing theory. Perceived content and creator cues constitute the antecedent evaluations, CSV represents the organism-level value appraisal, and purchase intention represents the response [19]. This mapping determines the structural ordering tested in the primary model.

SOR is complemented, rather than replaced, by additional theoretical perspectives. The Elaboration Likelihood Model distinguishes between informational and peripheral persuasion cues [20]; technology-acceptance models explain perceived usefulness and readiness to engage [10,11,21]; and uses-and-gratifications, social presence, and platform-affordance perspectives explain entertainment, interaction, and platform-mediated exposure [17,22,23].

This theoretical hierarchy avoids combining multiple frameworks without clearly defined roles. SOR specifies the primary sequence, whereas the other perspectives justify specific antecedents and contextual boundaries. Because the survey does not measure recommendation algorithms, actual exposure, or platform-level system interactions, these factors remain contextual rather than empirically tested mechanisms.

The resulting model is deliberately perception-based. It estimates associations among respondents' evaluations, CSV, and purchase intention, but does not establish the causal effects of changes in a video, creator, or platform feature.

2.3. Empirical Models and Research Gap

Table 1 summarizes selected empirical models linking SFV characteristics to trust, perceived value, attitude, immersive states, purchase intention, and observed purchase behavior. Most studies employ SEM or regression-based approaches, whereas relatively few integrate explanatory modeling with consumer segmentation, text analytics, and predictive modeling.

Table 1. Comparison of selected models of consumer responses and purchase-related outcomes in short-video marketing.

Reference	Context and Method	Main Mechanism	Statistically Significant Factors	Reported Evaluation/Outcome
Luo et al. [1]	Social media short videos; SEM; $N = 372$	Usefulness, ease of use, and entertainment → trust → purchase intention	Usefulness, ease of use, entertainment, trust	Short-video content characteristics significantly affect trust and purchase intention
Shen and Wang [2]	Short-video social commerce; SEM; $N = 350$	Persona perception → shared value creation → purchase intention	Credibility, willingness to use, consistency, completeness, clarity, likability, empathy, similarity; regulatory focus and social presence as moderators	$\chi^2/df = 2.173$; CFI = 0.911; TLI = 0.897; RMSEA = 0.058
Dwinanda et al. [3]	TikTok short-video ads; PLS-SEM; $N = 486$	Advertising features → advertising value and attitude → purchase intention	Personalization, entertainment, credibility, interactivity	Advertising value and attitude mediate the effects of TikTok ad features on purchase intention
Zhai et al. [4]	Video-based eWOM; SEM/regression-based testing	Product review video characteristics → purchase intention; involvement as moderator	Information quality, product visualization, emotional polarity, source credibility	Video-based eWOM affects purchase intention; involvement moderates selected relationships
Ngo et al. [6]	Generation Z consumers; SEM; $N = 350$	SFV marketing factors → brand attitude → purchase intention	Interesting content, perceived usefulness, scenario-based experience, interaction, enjoyment, celebrity involvement	$\chi^2/df = 1.718$; CFI = 0.951; TLI = 0.945; RMSEA = 0.045
Yin et al. [5]	Silver consumers on SFV platforms; SEM; $N = 284$	Social and technical platform characteristics → ERG needs → purchase intention	Social belonging, perceived trust, product relevance	R^2 for purchase intention = 0.460
Feng et al. [16]	Taobao mini-detail short videos; TAM2-based SEM; $N = 212$	Multimedia, virtual experience, recommendations → playfulness/usefulness → purchase intention	Multimedia effect, virtual experience, personalized recommendation, perceived playfulness, perceived usefulness	Perceived playfulness and usefulness mediate effects on purchase intention
Meng et al. [15]	TikTok SFV ads; grounded theory + OLS regression	Content characteristics → observed purchase behavior	Trustworthiness, expertise, attractiveness; nonlinear effects of authenticity and brand heritage	R^2 increases across model specifications; study links SFV ad content to observed behavior
Yin et al. [14]	SFV apps; PLS-SEM; $N = 398$	Platform affordances → desire to postpone closure → purchase intention	Product relevance, social interaction affordance, entertainment affordance, product visibility affordance	Adjusted R^2 : PI = 0.498; DPC = 0.419; PR = 0.473

Table 1. Cont.

Reference	Context and Method	Main Mechanism	Statistically Significant Factors	Reported Evaluation/Outcome
Yu et al. [18]	Furniture short-video ads; SOR-based SEM	Platform/content stimuli → flow and telepresence → purchase intention	Social influence, entertainment value, interactivity, facilitating conditions, media richness, flow, telepresence	Flow and telepresence significantly increase purchase intention
Yu and Wu [19]	Furniture-brand short videos; ELM + OLS regression	Execution features → communication effect	Live segments, graphics, subtitles/topics, music, title type, video duration, product category	Communication Effect Index used as performance proxy
Jiang and Chen [17]	SFV platforms; SEM	Use and gratification + social presence → perceived value and attitude → digital dependency → purchase intention	Gratification, social presence, perceived value, attitude, digital dependency	GFI = 0.925; AGFI = 0.911; NFI = 0.975; NNFI = 0.992; RMSEA = 0.0296
Present study	Bulgarian consumers; formative CSV PLS path model; N = 409	Clarity, likability, empathy, similarity, willingness to use → formative CSV → purchase intention	All five antecedents and CSV	R ² : CSV = 0.691; purchase intention = 0.322; CSV → PI β = 0.567

Note: The arrow (→) denotes a hypothesized directional relationship estimated in the structural model and does not, by itself, imply causality. All abbreviations used in the table are defined in the Abbreviations section.

Purchase intention remains the dominant outcome, mediated by constructs such as trust, advertising value, attitude, shared value, usefulness, flow, or digital dependency. Platform affordances and creator cues are increasingly prominent, but behavioral validation and European evidence remain limited.

The present study employs one primary structural model, in which CSV is specified formatively, and supplements it with complementary BI analyses. This analytical structure supports clear interpretation: PLS evaluates theory-driven associations, whereas clustering, sentiment analysis, and machine learning provide segmentation results, qualitative insights, and internal predictive evaluation, respectively.

2.4. Hypothesis Development

The hypotheses follow the SOR sequence defined above. Clarity, likability, empathy, similarity, and willingness to use are antecedent evaluations; CSV is a formative organism-level value appraisal; and purchase intention is the response. The construct is not treated as a synonym for trust or attitude [2,24–27].

2.4.1. Clarity

Clarity concerns whether viewers can understand and remember the content. In rapid scrolling environments, clear structure and unambiguous product information reduce processing effort. Prior SFV research associates understandability and ease of use with trust, value, and purchase-related responses [1,2]; clarity is therefore expected to support CSV.

2.4.2. Likability

Likability reflects whether the creator or presentation is perceived as pleasant, interesting, and socially appealing. Entertainment and creator appeal can strengthen attention and favorable evaluation [3,6,28], providing an affective basis for attributed value.

2.4.3. Empathy

Empathy reflects perceived understanding and emotional connection with the creator. Perspective-taking and relationship empathy can strengthen trust and behavioral response [29–32], suggesting a positive association with attributed product, social, and relational value.

2.4.4. Similarity

Similarity denotes perceived fit in interests, thinking, lifestyle, or values. Creator-viewer resemblance can increase identification, parasocial connection, and product interest [2,33], making the creator's activity more personally relevant and valuable.

2.4.5. Willingness to Use

Willingness to use reflects readiness to engage with creator content as an information or task resource. It is positioned before CSV because perceived usability creates opportunities to attribute value, whereas purchase intention concerns buying within six months. WTU4 mentions product decisions and could overlap conceptually with purchase intention; therefore, HTMT and a sensitivity model excluding WTU4 are reported. The ordering remains theoretical because the data are cross-sectional [10,11,34].

2.4.6. Creating Shared Value as a Formative Composite

CSV1–CSV8 cover distinct perceived contributions: product quality, relative sales, stakeholder interaction, employment, community welfare, local development, social improvement, and stable viewer relationships. These components need not move together and cannot be removed without narrowing the construct, so they form CSV rather than reflect one latent attitude [25–27,35].

CSV is explicitly perception-level and creator-specific. It represents value attributed by respondents, not audited evidence that objective Creating Shared Value has occurred [2,7].

2.4.7. Purchase Intention

Purchase intention represents a stated likelihood of buying and is a proximal but imperfect proxy for behavior [8,9]. In SFV research, it is associated with content, trust, entertainment, usefulness, social presence, creator perceptions, and value appraisals [1,2,6,16,17].

Based on the synthesis of previous studies and the proposed conceptual model, the research hypotheses are formulated as follows:

H1: *Clarity is positively associated with CSV in short-video marketing.*

H2: *Likability is positively associated with CSV in short-video marketing.*

H3: *Empathy is positively associated with CSV in short-video marketing.*

H4: *Similarity is positively associated with CSV in short-video marketing.*

H5: *Willingness to use is positively associated with CSV in short-video marketing.*

H6: *CSV is positively associated with purchase intention.*

H7: *Demographic and usage classifications are associated with between-group differences in the CSV indicator mean and/or purchase-intention mean.*

H1–H6 specify the structural associations between the five antecedents and CSV and between CSV and purchase intention, whereas H7 addresses differences across demographic and usage groups (Figure 1).

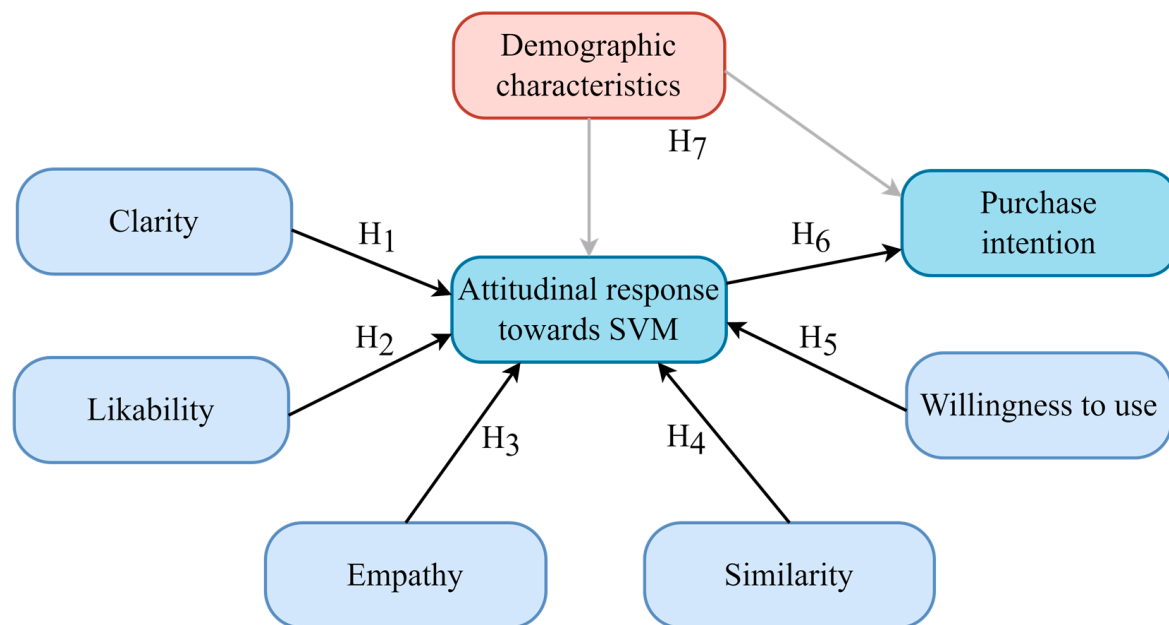


Figure 1. Conceptual model of the hypothesized structural associations and the demographic and usage-group differences addressed by H7.

3. Research Methodology

The objective is to examine associations between consumers' perceptions of commercial SFV, CSV, and purchase intention and to assess demographic and usage-group differences. The cross-sectional survey supports standardized measurement but does not identify causal effects [36,37]. The empirical strategy distinguishes the primary explanatory model from complementary BI procedures for description, segmentation, text interpretation, group-difference assessment, and prediction [12].

The BI-oriented approach integrates structured and unstructured consumer-response data. Descriptive profiling summarizes the sample; clustering identifies response patterns; sentiment analysis interprets open-ended feedback; formative PLS estimates construct associations; one-way ANOVA assesses demographic and usage-group differences; and machine learning evaluates internal predictive performance.

3.1. Questionnaire Design and Data Collection

Primary data were collected through a self-administered online questionnaire. This format was selected because it enables efficient distribution, standardized measurement, respondent anonymity, and rapid data collection from digitally active consumers, which is suitable for a study focused on short-video and social-media behavior [36]. The questionnaire was developed using established principles of survey and scale design, including clear item wording, logically grouped question blocks, and multi-item measurement of latent constructs [36,37].

The questionnaire focused on consumer responses to short videos on social media platforms in Bulgaria. The introductory text explained that the study concerned commercial short-video content, including beauty and makeup tutorials, fashion and styling, cooking and recipes, educational content, life hacks and advice, fitness and workouts, pet and animal videos, travel and adventure, transformation videos, motivational and inspirational content, gaming-related material, unboxing, and reviews. Respondents were informed that the survey was anonymous, that completion would take approximately five minutes, and that participants would receive access to summary results.

The instrument contained demographic and usage questions, multi-item perception measures, shared-value and purchase-intention items, and open-ended feedback. The survey link was distributed through partner organizations and institutional websites, email, and social media. Participation was voluntary and self-selected; no sampling frame or invitation denominator was available, so a response rate and formal nonresponse-bias estimate could not be calculated. The final sample contains 409 valid responses.

The primary PLS model uses 28 indicators because PI2 is excluded from the purchase-intention construct, whereas all 29 indicators are retained in the complementary analyses and in the sensitivity analysis involving PI2.

The measurement items were adapted primarily from the short-video social-commerce framework developed by Shen and Wang [2], with the item statements adjusted to fit the context of the present study. The questionnaire covered credibility, consistency, completeness, clarity, likability, empathy, similarity, willingness to use, CSV, and purchase intention. The empirical analyses focus on 29 indicators from Q14–Q20 measuring clarity, likability, empathy, similarity, willingness to use, CSV, and purchase intention.

3.2. Questionnaire Measurements and Scales

The questionnaire comprised 22 questions. Questions Q1–Q10 collected demographic, socio-economic, and behavioral information. These items covered gender, age group, place of residence, municipality, average monthly income per household member, educational level, social-media experience, frequency of social-media use, number of followed influencers, and daily time spent watching short videos.

Questions Q11–Q20 were implemented as multi-item grid questions using a five-point Likert-type response scale, ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). Likert-type scales are widely used for measuring attitudes and perceptions because they allow respondents to express the degree of agreement with a set of statements representing an underlying construct [38]. The use of multiple indicators per construct also supports reliability and validity assessment in SEM-based analysis [13,37].

Table 2 consolidates the constructs used in the final model, their item content and sources, measurement direction, and principal estimates. All items used a five-point agreement scale. Item codes are retained here because this is the measurement section.

Table 2. Constructs, item content, sources, measurement specification, and estimates.

Construct and Codes	Abbreviated Item Content and Source
Clarity CLA1–CLA3	Content clear enough to view; information easy to understand; creator memorable. Adapted from Chu et al. [39].
Likability LIK1–LIK4	Likable; friendship affinity; time-spending affinity; interesting. Adapted from Borner and Brangier [40].
Empathy EMP1–EMP3	Understanding; felt ties; imagined daily life. Adapted from Mangus et al. [32].
Similarity SIM1–SIM4	Self-similarity; thinking alike; shared interests; likely agreement. Adapted from Xiao et al. [41].
Willingness to use WTU1–WTU4	Use in a posting task; learn more; use in a video task; support product decisions. Adapted from Aparicio et al. [34].
Creating Shared Value CSV1–CSV8	Product quality; relative sales; stakeholder interaction; jobs; community welfare; local development; social improvement; stable relationships. Adapted from Ham et al. [35].
Purchase intention PI1–PI3	Purchase probability and likelihood within six months. PI1 and PI3 in the primary PLS model; PI2 retained for sensitivity and complementary analyses. Adapted from Kim et al. [42].

3.3. BI-Oriented Data Analysis Methods

The analyses were organized as complementary BI procedures with distinct evidential purposes: description, structural association testing, segmentation, textual interpretation, group-difference assessment, and internal prediction.

First, the dataset was examined for completeness, consistency, and plausibility. Closed-ended questions were encoded numerically according to the predefined questionnaire coding scheme. Open-ended responses were processed separately, including text normalization and preparation for qualitative and sentiment analysis.

Second, descriptive statistics summarized the sample structure and short-video usage behavior. Frequency distributions and percentages were used for gender, age, residence, education, income, social-media experience, use frequency, followed-influencer count, and viewing time. These statistics provide the baseline context for the subsequent modeling results [43].

Third, PLS path modeling tested the six structural hypotheses. The model was estimated in SmartPLS 3.2.9 (SmartPLS GmbH, Bönningstedt, Germany) [44] using the path-weighting scheme, standardized indicators, Mode A for the common-factor constructs, and Mode B for CSV. The estimation used a maximum of 1000 iterations, a convergence tolerance of 10^{-9} , complete-case data, and 10,000 bootstrap samples with two-tailed percentile confidence intervals. Mode A constructs were assessed using indicator loadings, Cronbach's alpha, composite reliability (CR), average variance extracted (AVE), and HTMT. CSV was assessed using content coverage, indicator VIFs, outer weights, absolute loadings, and its structural relationships [13,25–27,45].

H7 was tested using one-way ANOVA in IBM SPSS Statistics 27.0.0.0 (IBM Corp., Armonk, NY, USA). Each of the nine demographic and usage classifications was examined separately for two observed outcomes: the equal-weight mean of CSV1–CSV8, referred to as the CSV indicator mean, and the mean of PI1–PI3, referred to as the purchase-intention indicator mean. These observed scale means were used specifically for the group-comparison analyses and were distinct from the Mode B CSV composite score and the two-item purchase-intention construct estimated in the principal PLS model. The original questionnaire categories were retained in the omnibus analyses. Eta squared (η^2) was used to quantify the omnibus effect size, and Holm's procedure controlled the family-wise error rate across the family of 18 ANOVA tests [43,46].

PLS-based interaction analysis and multigroup analysis, together with measurement-invariance assessment, represent alternative approaches when the objective is to determine whether structural-path coefficients vary across groups [47–50]. These procedures were not applied because H7 concerns omnibus differences in observed outcome means across the original demographic and usage categories rather than moderation of the relationships in the PLS model. One-way ANOVA was therefore preferred because it directly addresses H7, accommodates grouping variables with more than two categories without arbitrary dichotomization, and maintains a clear distinction between the group-comparison analysis and the principal PLS model.

Fourth, hierarchical and k-means clustering were used as exploratory BI segmentation methods. The elbow method and silhouette coefficient informed the comparison of two- to five-cluster solutions [51,52]. Because all perception and outcome indicators enter the clustering, the segments are interpreted as general response profiles rather than validated market personas.

Fifth, an illustrative BI responsiveness score combined construct averages using normalized antecedent coefficients. It ranks anonymous response profiles and is presented as transparent descriptive scoring, not a formal MCDM application, causal conversion rule, or validated campaign-optimization method [53].

Sixth, machine-learning models supplied predictive BI evidence by estimating the equal-weight mean of CSV1–CSV8 from antecedent indicators using Orange Data Mining 3.22.0. A 70:30 train–test split and training-set cross-validation were used; MSE, RMSE, MAE, and R^2 evaluated internal test-set performance [12,54].

Finally, sentiment analysis supplied text-based BI evidence from open-ended responses (Python 3.11.2). Text normalization, translated polarity scoring, and manual thematic interpretation identified recurring views concerning usefulness, entertainment, authenticity, credibility, relatability, disclosure, and purchase motivation [55].

The questionnaire, response dataset, and supporting documentation are made available through Mendeley Data [56].

4. Results

Descriptive, segmentation, and sentiment results are followed by the formative-CSV structural model, the ANOVA test of H7, and predictive BI analysis.

4.1. Data Preparation and Sample Description

The survey link was distributed through the authors' institutional websites, email channels, partner organizations, and social media platforms. Participation was voluntary and anonymous. Data collection targeted Bulgarian online consumers and was conducted between 11 January 2025 and 25 February 2025. After removing incomplete entries, the final analytical sample comprised 409 valid responses. The questionnaire, response dataset, and supporting documentation are publicly available through Mendeley Data [56].

All closed-ended questions (Q1–Q3 and Q5–Q20) were encoded using predefined coding rules. The open-ended fields—municipality, opinions/suggestions, and followed/favorite short-video influencers—were processed separately through spelling normalization, text cleaning, and preparation for qualitative and sentiment analysis. Duplicate screening was performed on the full coded response dataset. Repeated response profiles in the 29-item perception matrix were retained because identical Likert-scale patterns can occur naturally in ordinal survey data and do not by themselves indicate invalid responses.

To visualize similarity among respondents, an ordered Euclidean dissimilarity matrix was computed using the 29 standardized indicators representing clarity, likability, empathy, similarity, willingness to use, CSV, and purchase intention (Figure 2). Rows and columns represent respondents; lighter values indicate more similar response profiles, whereas darker cyan values indicate larger Euclidean distances. The matrix was visualized using the `fviz_dist()` function from the `factoextra` package in R 4.4.1 (R Core Team, R Foundation for Statistical Computing, Vienna, Austria) [57,58].

To profile the respondents and summarize key patterns, descriptive statistics and distribution analyses were conducted. Table 3 reports the main demographic and usage characteristics. Most respondents were female (70.7%), and 76.6% were aged 30 or younger. The sample was predominantly urban and geographically concentrated in Bulgaria's South Central Region.

In terms of geographic distribution, municipality names were normalized to correct spelling and capitalization variants before calculating regional distributions. The largest share of respondents was from the Plovdiv municipality (204 respondents; 49.9%), followed by Haskovo (19 respondents; 4.6%) and Pazardzhik (13 respondents; 3.2%). Overall, the sample was strongly concentrated in the South Central Region (approximately 350 respondents; 85.6%), followed by the South East Region (30 respondents; 7.3%) and the South West Region (17 respondents; 4.2%).

The high relevance of social-media and short-video consumption in the sample is consistent with broader Bulgarian digital-use patterns. DataReportal reports that Bulgaria

had 5.86 million Internet users at the start of 2025, with 87.1% Internet penetration, and 4.37 million social-media user identities, equivalent to 64.9% of the population. It also reports YouTube advertising reach of 4.37 million users and TikTok advertising reach of 2.28 million adults in Bulgaria in early 2025 [59].

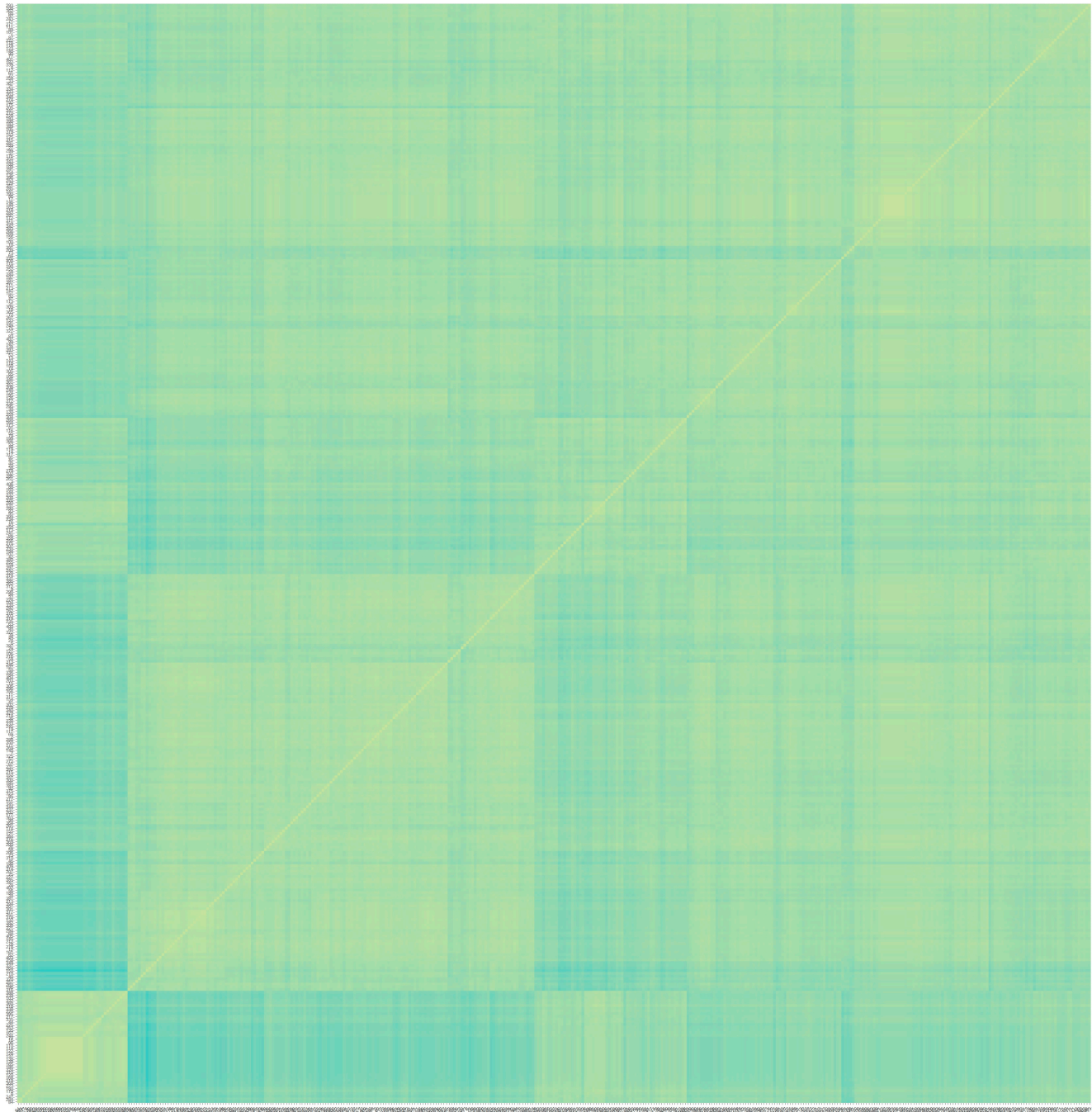


Figure 2. Ordered dissimilarity matrix based on the 29 standardized perception and outcome indicators. Note: Some labels overlap because Figure 2 displays all 409 observations on both the x- and y-axes. For further details, please refer to the dataset available in the Mendeley Data repository [56].

Table 3. Respondent profile and short-video usage characteristics.

Variables of the Sample		No. of Respondents	Percentage (%)
1. Gender	Male	120	29.3
	Female	289	70.7
2. Age	Under 20	186	45.5
	Between 21 and 30	127	31.1
	Between 31 and 40	32	7.8
	Between 41 and 50	35	8.6
	Over 50	29	7.1
3. Place of residence	City/over 50,000 inhabitants	236	57.7
	Town/1000–50,000 inhabitants	161	39.4
	Village/under 1000 inhabitants	12	2.9
4. Municipality		-	-
5. Monthly income per household member	Less than BGN 1454	170	41.6
	More than BGN 1454	239	58.4
6. Education	High school	277	67.7
	Bachelor	75	18.3
	Master	49	12.0
	PhD	8	2.0
7. Experience with social media	Less than 3 years	23	5.6
	3 to 5 years	54	13.2
	More than 5 years	332	81.2
8. Frequency of use of social media	Less than once a week	4	1.0
	Once or twice a week	6	1.5
	Several times a week	9	2.2
	Once or twice a day	38	9.3
	Several times a day	233	57.0
	Several times an hour	119	29.1
9. Number of influencers that you follow on social media	Less than 10	188	46.0
	10 to 20	103	25.2
	20 to 30	53	13.0
	More than 30	65	15.8
10. Time per day spent watching short videos on social media	Less than an hour	150	36.7
	Between 1 and 2 h	140	34.2
	Between 2 and 3 h	76	18.6
	More than 3 h	43	10.5

Most respondents reported very frequent social-media use: 95.4% accessed social media at least once per day, including 57.0% several times per day and 29.1% several times per hour. Exposure to creator content was also substantial, with 54.0% following 10 or more influencers. Short-video consumption was similarly intensive: 63.3% watched short videos for one hour or more per day. These patterns confirm the relevance of SFV marketing for this digitally active, comparatively young sample.

4.2. Exploratory BI Segmentation Analysis

To illustrate heterogeneity in respondents' perceptions, hierarchical clustering was applied to the same 29 standardized indicators. Figure 3 presents clustering across respondents, while Figure 4 presents clustering across indicators. Cell colors range from the lowest standardized values (white; approximately -2.31) through pale yellow-green tones to the highest values (blue; approximately 2.75) (Figure 3). The respondent dendrogram summarizes similar response profiles, whereas the indicator dendrogram shows variables that tend to co-vary. The heat maps and dendrograms were generated with the Heat Map widget in Orange Data Mining 3.22.0 (Bioinformatics Laboratory, University of Ljubljana, Ljubljana, Slovenia) [60].

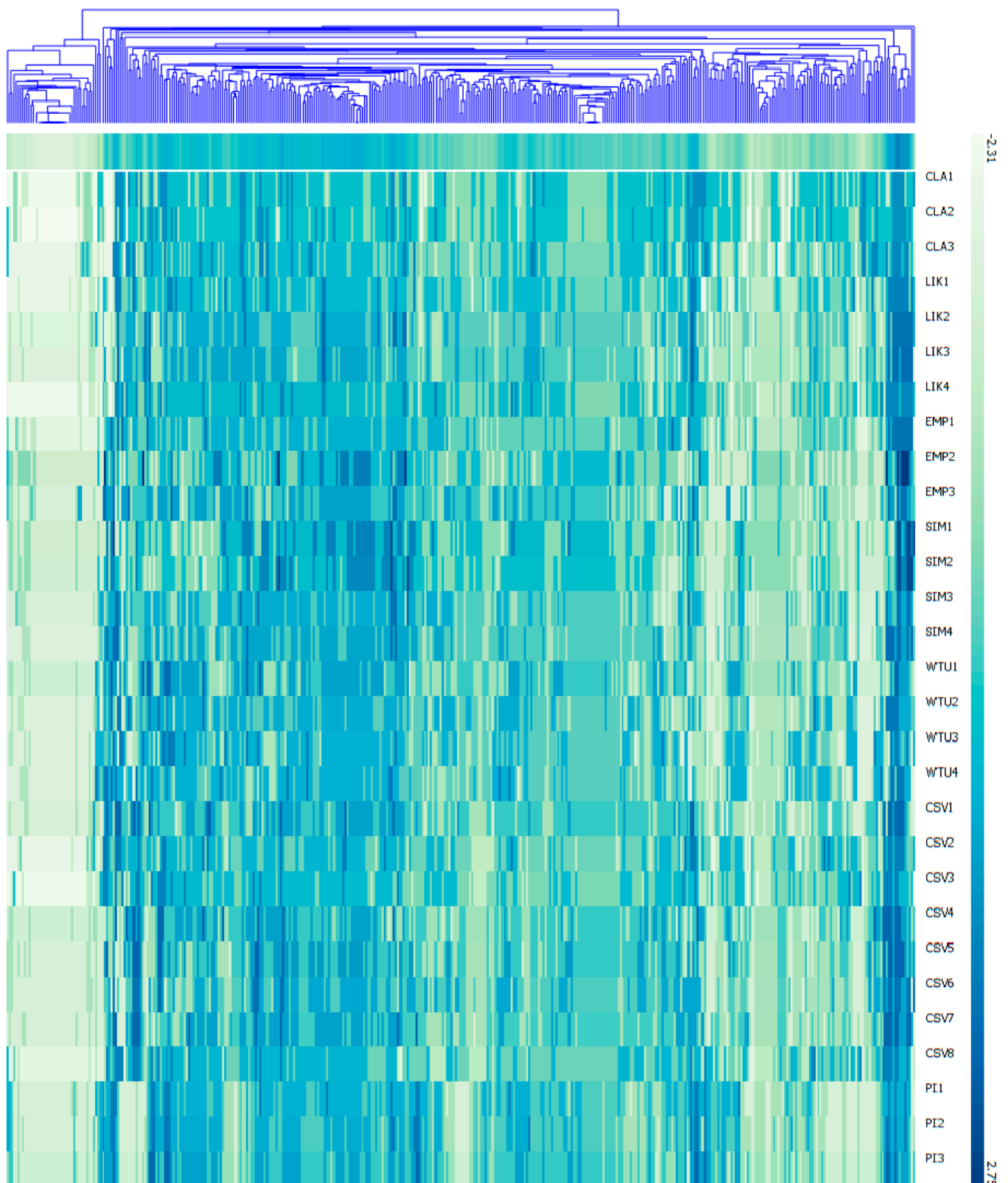


Figure 3. Hierarchical heat map of respondents using the 29 standardized perception and outcome indicators.

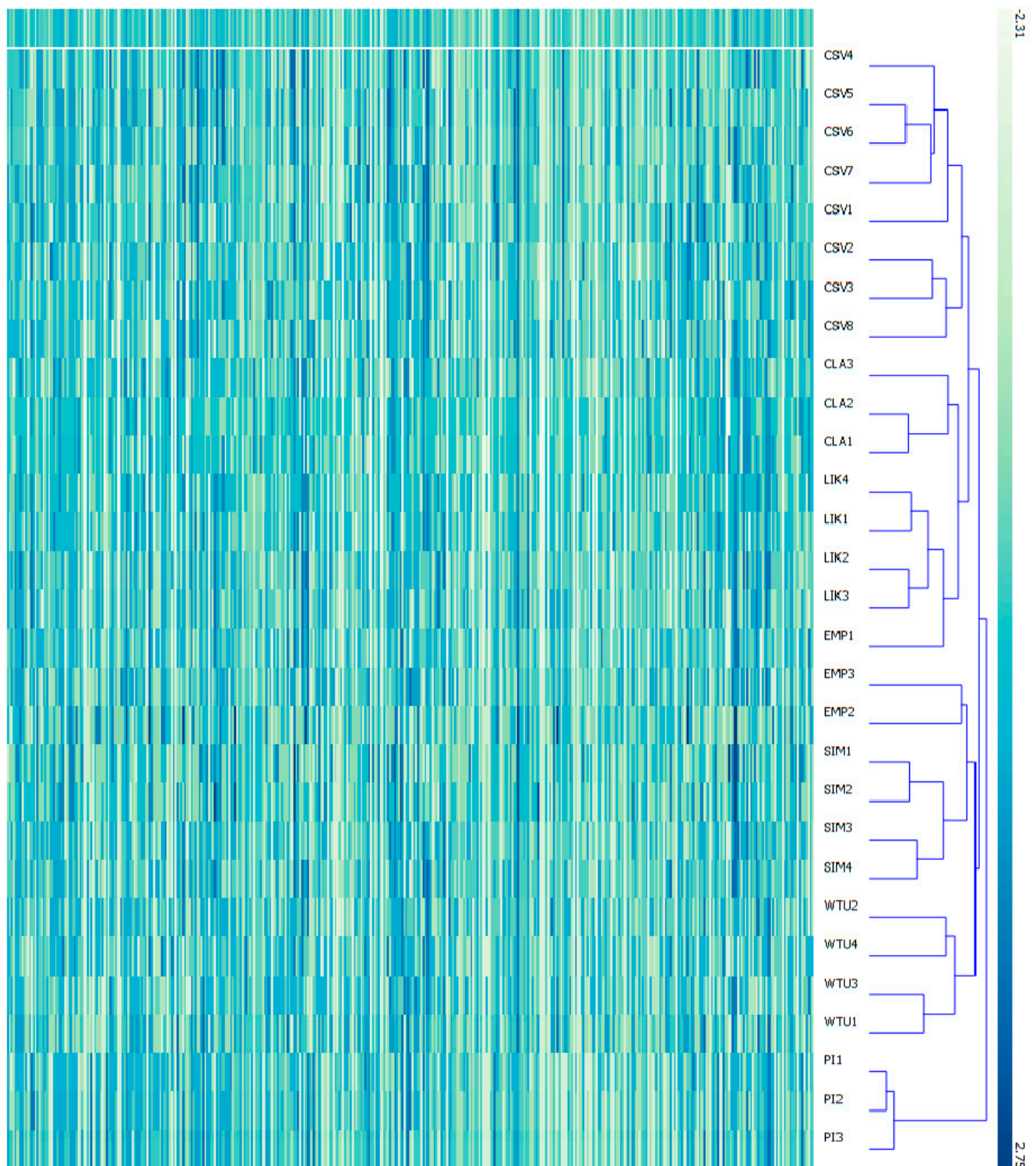


Figure 4. Hierarchical heat map of the 29 standardized perception and outcome indicators.

To segment consumers based on similar perception profiles, k-means clustering was applied to the same 29 standardized indicators. The number of clusters was examined for $k = 2, 3, 4$, and 5 . The elbow method and silhouette coefficient indicated that the two-cluster solution was the most interpretable and statistically preferable [32,33]. The two clusters comprised 309 and 100 respondents, respectively. As illustrated in Figure 5, the $k = 2$ solution separates a large favorable segment from a smaller less favorable segment.

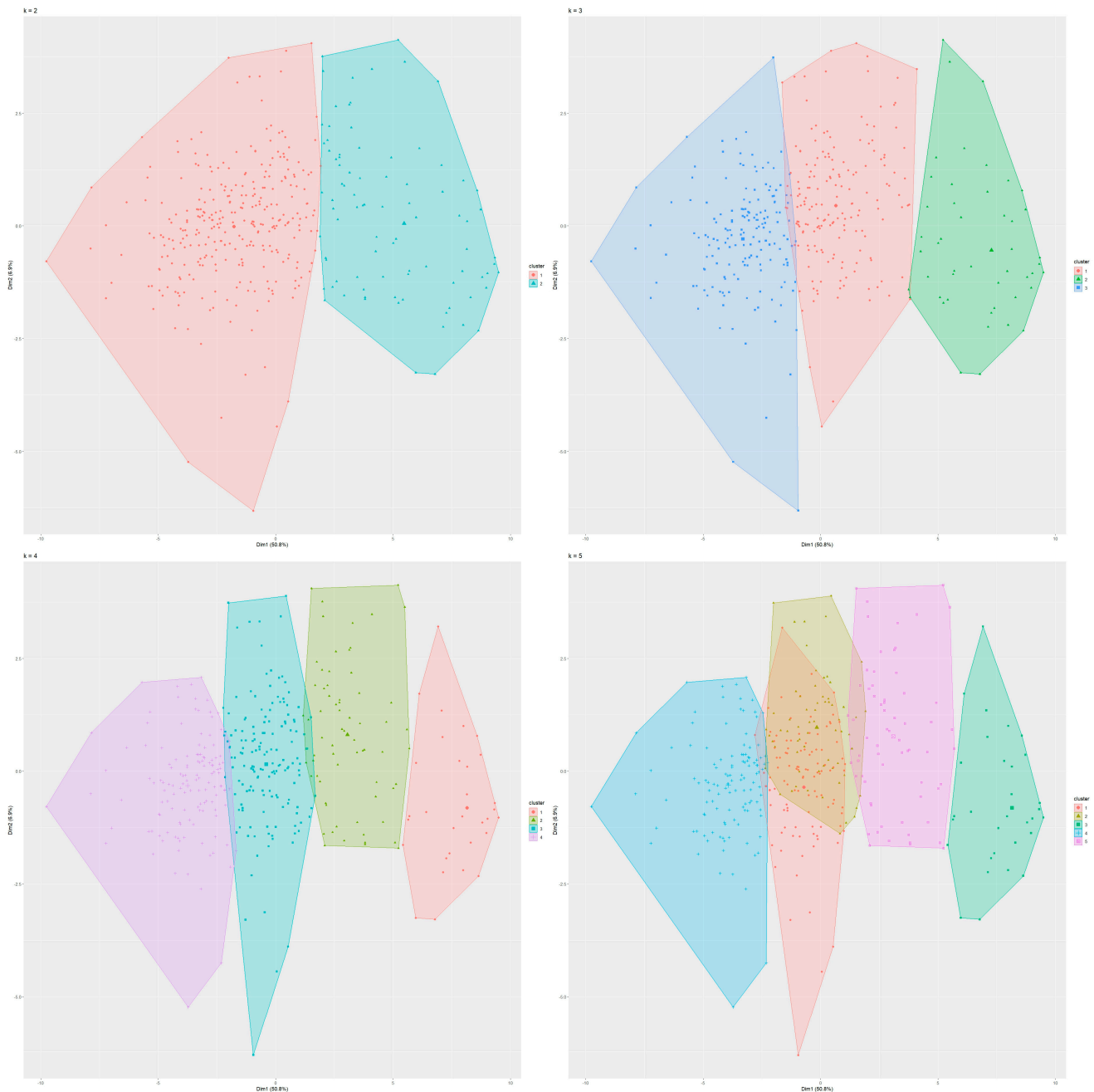


Figure 5. K-means solutions for $k = 2, 3, 4$, and 5 using the 29 standardized perception and outcome indicators.

Cluster 1 represents a generally favorable response profile, whereas Cluster 2 represents a generally less favorable profile (Table 4). The largest differences concern similarity, likability, and willingness to use. Because the same response matrix contains antecedents, CSV items, and purchase intention, the result may partly reflect general positivity or response style and is not interpreted as a stable market taxonomy.

A demographic comparison of the two clusters shows that Cluster 1 is more strongly represented by younger respondents. In Cluster 1, 80.3% of respondents are aged 30 or younger, compared with 65.0% in Cluster 2. Cluster 1 also contains a higher share of respondents following 10 or more influencers (62.8% vs. 27.0%) and watching short videos for at least one hour per day (67.3% vs. 51.0%). Chi-square tests indicate statistically significant differences between clusters for age, education, number of followed influencers,

and time spent watching short videos, whereas gender, residence, income, social-media experience, and frequency of use do not show significant differences.

Table 4. Mean indicator values by exploratory response-profile cluster.

	CLA1	CLA2	CLA3	LIK1	LIK2	LIK3	LIK4	EMP1	EMP2	EMP3	SIM1	SIM2
Cluster 1	3.728	3.851	3.573	3.621	3.346	3.275	3.777	3.440	2.595	3.142	2.722	2.754
Cluster 2	2.510	2.530	2.230	2.050	1.810	1.650	2.070	1.830	1.550	1.610	1.450	1.440
Difference	1.218	1.321	1.343	1.571	1.536	1.625	1.707	1.610	1.045	1.532	1.272	1.314
	SIM3	SIM4	WTU1	WTU2	WTU3	WTU4	CSV1	CSV2	CSV3	CSV4	CSV5	CSV6
Cluster 1	3.304	3.275	3.052	3.382	3.353	3.356	3.129	3.505	3.618	3.061	3.078	3.236
Cluster 2	1.590	1.600	1.620	1.670	1.830	1.850	1.610	2.110	2.230	1.610	1.670	1.640
Difference	1.714	1.675	1.432	1.712	1.523	1.506	1.519	1.395	1.388	1.451	1.408	1.596
	CSV7	CSV8	PI1	PI2	PI3							
Cluster 1	3.149	3.508	3.269	3.181	3.019							
Cluster 2	1.680	2.090	1.830	1.820	1.690							
Difference	1.469	1.418	1.439	1.361	1.329							

4.3. Exploratory Text and Sentiment Analysis

The open-ended comments were originally written mainly in Bulgarian. Before sentiment analysis, they were translated into English and manually checked for obvious mistranslations and polarity errors. TextBlob polarity scores were then used as exploratory sentiment indicators, supplemented by manual thematic interpretation. Because the sentiment classification relies on translated open-ended responses and a lexicon-based English-language tool, the polarity results are interpreted as exploratory and are used mainly to support the qualitative thematic interpretation.

The first open-ended item contained 237 non-empty entries. After removing blank, duplicated, and non-informative responses, 181 usable comments were retained for sentiment analysis. A lexicon-based approach was applied using TextBlob version 0.15.3 with PatternAnalyzer in Python version 3.11.2 (Python Software Foundation, Wilmington, DE, USA) with polarity scores ranging from -1 to $+1$ and a three-class scheme: positive (>0.05), neutral (-0.05 to 0.05), and negative (<-0.05) [61]. Positive comments ($n = 131$; mean polarity = 0.281) emphasized rapid attention capture, effective product demonstration, discovery, and support for purchase decisions. Neutral comments ($n = 39$; mean polarity = 0.004) were conditional on sufficient product detail and disclosure of paid partnerships. Negative comments ($n = 11$; mean polarity = -0.328) highlighted distraction, information overload, and distrust, reinforcing the need for transparency, authenticity, and credibility.

The second open-ended item contained 247 non-empty entries, of which 168 remained after cleaning. Respondents preferred authentic, experience-based content from creators who had tested products, reported drawbacks, and avoided overly forced promotion. Clear, concise demonstrations and practical information were valued, as were relatable everyday stories, humor, and entertainment. A smaller group emphasized sustainability and healthy living. Overall, the comments favor content that combines entertainment with credible and useful information aligned with viewers' interests.

The same responses also reveal a reach-based visibility effect. The most frequently mentioned influencers were also among highly visible Bulgarian creators on Instagram and TikTok, suggesting that respondent nominations partly reflect platform visibility and perceived reach. Because follower counts change frequently, exact follower numbers are not reported in the final manuscript.

4.4. PLS-SEM Model with Formative Creating Shared Value

PLS-SEM was used to estimate the associations among consumers' perceptions of short-video content and creators, CSV, and purchase intention. The reanalysis used all

409 complete observations. Clarity, likability, empathy, similarity, willingness to use, and purchase intention were specified reflectively (Mode A). CSV was specified formatively (Mode B) because CSV1–CSV8 represent distinct product/economic, community/social, and relational contributions that jointly constitute perceived shared value and are not interchangeable manifestations of one latent attitude [24–27]. The label CSV is used because the questionnaire captures value attributed by respondents rather than verified economic or social outcomes.

4.4.1. Model Specification and Estimator Verification

The structural specification retained the six hypothesized paths: clarity, likability, empathy, similarity, and willingness to use were modeled as antecedents of CSV, and CSV was modeled as an antecedent of purchase intention. PI1 and PI3 were retained for purchase intention, while PI2 remained excluded because of strong semantic redundancy with the other two items. Indicators were standardized, the path-weighting scheme was used, and significance was assessed with 10,000 bootstrap samples, two-tailed tests, and percentile confidence intervals. The formative model was independently re-estimated using a reproducible path-weighting PLS-PM implementation. As an implementation check, re-estimating the original all-reflective model reproduced every SmartPLS coefficient reported in the submitted version to three decimal places.

4.4.2. Reflective Measurement-Model Assessment

The reflective constructs were evaluated using indicator loadings, Cronbach’s alpha, composite reliability (CR), average variance extracted (AVE), and the heterotrait–monotrait ratio (HTMT). All loadings exceeded 0.81, alpha ranged from 0.802 to 0.919, CR ranged from 0.883 to 0.960, and AVE ranged from 0.705 to 0.923 (Table 5). These results satisfy conventional thresholds of 0.70 for loadings and reliability and 0.50 for AVE. HTMT was evaluated only among reflective constructs because it is not an appropriate validity criterion for formative CSV; the largest value was 0.845 for empathy and similarity [13,62–64].

Table 5. Reflective measurement-model assessment.

Construct	Items	Loading Range	Alpha	CR	AVE
Clarity	3	0.842–0.920	0.862	0.916	0.785
Likability	4	0.884–0.913	0.919	0.943	0.804
Empathy	3	0.812–0.888	0.802	0.883	0.717
Similarity	4	0.845–0.894	0.895	0.927	0.762
Willingness to use	4	0.834–0.853	0.861	0.905	0.705
Purchase intention	2	0.960–0.961	0.914	0.960	0.923

Note: Alpha, CR, and AVE apply only to the reflective constructs.

4.4.3. Formative CSV Assessment

Formative CSV was assessed using indicator collinearity, the significance and relevance of outer weights, and absolute outer loadings. Indicator VIF values ranged from 2.285 to 3.969; none exceeded the applied threshold of 5, although CSV5 and CSV6 exceeded the more conservative value of 3 (Table 6). Six outer weights were positive and statistically significant. CSV4 was positive but nonsignificant, whereas CSV5 had a small negative nonsignificant weight. Both retained substantial positive loadings (0.728 and 0.754), indicating absolute relevance despite weak unique contributions after controlling for the other indicators. They were retained to preserve employment and community-welfare content. Internal-consistency reliability and AVE were not calculated for CSV because those statistics assume reflective measurement. Redundancy validity could not be tested because the questionnaire did not include an independent global CSV item; this is a limitation [26,27].

Table 6. Formative CSV indicator assessment.

Indicator	Weight	SE	<i>p</i>	95% CI	Loading	VIF
CSV1	0.195	0.059	<0.001	[0.081, 0.311]	0.815	2.358
CSV2	0.178	0.052	<0.001	[0.075, 0.279]	0.779	2.296
CSV3	0.226	0.061	<0.001	[0.105, 0.346]	0.811	2.404
CSV4	0.073	0.053	0.168	[−0.034, 0.176]	0.728	2.408
CSV5	−0.064	0.074	0.389	[−0.213, 0.078]	0.754	3.969
CSV6	0.211	0.077	0.006	[0.061, 0.361]	0.799	3.506
CSV7	0.254	0.060	<0.001	[0.132, 0.368]	0.843	2.737
CSV8	0.162	0.062	0.009	[0.041, 0.289]	0.811	2.285

Note: Weights express unique relative contributions; loadings express absolute indicator-CSV associations. A nonsignificant weight does not by itself require indicator deletion when loading relevance and content coverage support retention.

4.4.4. Structural-Model Assessment

All six hypothesized structural associations were positive and statistically significant. Clarity showed the largest antecedent coefficient, followed by willingness to use, similarity, likability, and empathy. The five antecedents explained 69.1% of CSV variance. CSV was positively associated with purchase intention ($\beta = 0.567$, 95% CI [0.495, 0.646]) and explained 32.2% of its variance (Table 7 and Figure 6). Structural VIF values ranged from 2.120 to 3.635, remaining below 5.

Table 7. Structural paths for the formative-CSV model.

Relationship	β	SE	<i>t</i>	<i>p</i>	95% CI	<i>f</i> ²
Clarity → CSV	0.260	0.052	4.996	<0.001	[0.166, 0.369]	0.102
Likability → CSV	0.177	0.069	2.549	0.011	[0.036, 0.307]	0.026
Empathy → CSV	0.130	0.051	2.546	0.011	[0.035, 0.231]	0.019
Similarity → CSV	0.205	0.046	4.435	<0.001	[0.114, 0.295]	0.044
Willingness to use → CSV	0.226	0.053	4.255	<0.001	[0.120, 0.328]	0.076
CSV → Purchase intention	0.567	0.039	14.683	<0.001	[0.495, 0.646]	0.474

Note: The arrow (→) denotes a hypothesized directional relationship estimated in the structural model and does not, by itself, imply causality.

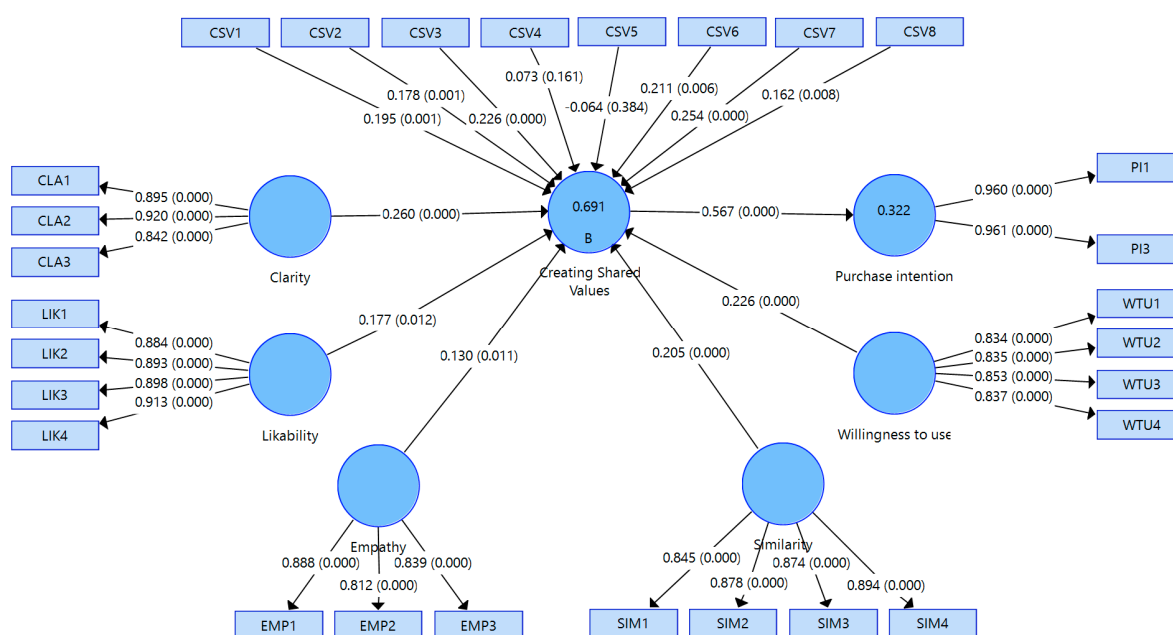


Figure 6. Formative-CSV PLS-SEM model with outer weights, path coefficients, and coefficients of determination. CSV4 and CSV5 have nonsignificant outer weights; two-tailed inference uses 10,000 bootstrap samples.

Out-of-sample predictive performance was assessed using repeated 10-fold cross-validation. The indicator-level Q^2_{predict} values were 0.278 for PI1 and 0.291 for PI3, with corresponding PLS-SEM RMSE values of 1.048 and 0.959. Both Q^2_{predict} values were positive, indicating that the model outperformed the naïve training-sample-mean benchmark for both purchase-intention indicators [65,66].

4.4.5. Robustness and Specification Assessment

The principal structural result remained stable when PI2 was restored and when WTU4 was removed. With all three purchase-intention indicators, the CSV–purchase-intention coefficient was 0.564, and the R^2 for purchase intention was 0.318. Excluding WTU4 reduced the willingness-to-use–CSV coefficient from 0.226 to 0.193, while the CSV–purchase-intention coefficient remained 0.567. Thus, the principal CSV association does not appear to depend materially on the exclusion of PI2 or the inclusion of WTU4.

Alternative structural specifications did not support a strict full-mediation interpretation. Adding a direct willingness-to-use–purchase-intention path produced significant coefficients for CSV ($\beta = 0.369$, $p < 0.001$) and willingness to use ($\beta = 0.285$, $p < 0.001$), increasing the R^2 for purchase intention to 0.364. A broader exploratory model with direct antecedent paths retained a significant CSV path ($\beta = 0.255$, $p = 0.003$) and significant direct associations for empathy and willingness to use, increasing R^2 to 0.393. A two-dimensional robustness model also identified positive associations between purchase intention and both economic contribution ($\beta = 0.231$, $p = 0.003$) and social contribution ($\beta = 0.364$, $p < 0.001$).

The CSV–purchase-intention association remained positive in the sensitivity and direct-path specifications, while both CSV dimensions were positively associated with purchase intention in the two-dimensional model. Direct antecedent associations may coexist with the CSV pathway; therefore, these analyses do not independently establish mediation, complete mediation, or a unique causal mechanism.

4.5. Demographic and Usage Group Differences

H7 was evaluated using the original questionnaire categories for all nine demographic and usage classifications. Table 8 reports all 18 omnibus tests, together with their η^2 effect sizes, raw p -values, and Holm-adjusted p -values. The CSV outcome is the equal-weight mean of CSV1–CSV8, whereas the purchase-intention outcome is the mean of PI1–PI3. These observed composite scores are used for the ANOVA analyses and are distinct from the latent-variable scores estimated in the primary PLS model.

Table 8. One-way ANOVA results for H7 across the CSV indicator-mean and purchase-intention mean.

Grouping Variable	CSV Indicator Mean: F(df1,df2); η^2	CSV p /pHolm	Purchase Intention: F(df1,df2); η^2	PI p /pHolm	H7 After Holm
Gender	0.083(1,407); 0.000	0.773/1.000	4.161(1,407); 0.010	0.042/0.420	No
Age	5.416(4,404); 0.051	<0.001/0.004	6.234(4,404); 0.058	<0.001/0.001	Both outcomes
Residence size	0.331(2,406); 0.002	0.718/1.000	0.671(2,406); 0.003	0.512/1.000	No
Household income	0.570(1,407); 0.001	0.451/1.000	1.163(1,407); 0.003	0.281/1.000	No
Education	3.464(3,405); 0.025	0.016/0.180	2.708(3,405); 0.020	0.045/0.420	No
SM experience	0.480(2,406); 0.002	0.619/1.000	0.901(2,406); 0.004	0.407/1.000	No
SM use frequency	2.085(5,403); 0.025	0.066/0.531	4.295(5,403); 0.051	<0.001/0.011	PI
Followed influencers	14.452(3,405); 0.097	<0.001/<0.001	10.473(3,405); 0.072	<0.001/<0.001	Both outcomes
Daily SFV time	4.402(3,405); 0.032	0.005/0.055	8.957(3,405); 0.062	<0.001/<0.001	PI

After Holm correction across the family of 18 ANOVAs, age showed significant differences in both the CSV indicator mean ($F(4,404) = 5.416$, $p_{\text{Holm}} = 0.004$, $\eta^2 = 0.051$) and the purchase-intention mean ($F(4,404) = 6.234$, $p_{\text{Holm}} = 0.001$, $\eta^2 = 0.058$). The number of followed influencers also showed significant differences in both outcomes (CSV: $F(3,405) = 14.452$, $p_{\text{Holm}} < 0.001$, $\eta^2 = 0.097$; purchase intention: $F(3,405) = 10.473$, $p_{\text{Holm}} < 0.001$, $\eta^2 = 0.072$). Social-media use frequency ($F(5,403) = 4.295$, $p_{\text{Holm}} = 0.011$,

$\eta^2 = 0.051$) and daily SFV viewing time ($F(3,405) = 8.957$, $p_{\text{Holm}} < 0.001$, $\eta^2 = 0.062$) showed significant differences only in purchase intention. Descriptively, respondents following fewer than 10 influencers had lower mean scores for both outcomes than the other follower-count groups. Purchase-intention means were generally higher among respondents using social media several times per day or hour and among those viewing SFVs for 1–3 h per day; however, these patterns were not monotonic, and the smallest social-media-use categories contained only 4–9 respondents. Gender, residence size, household income, education, and social-media experience did not distinguish either outcome after correction. H7 is therefore partially supported at the omnibus mean-difference level.

4.6. Predictive BI Analysis and Illustrative Responsiveness Scoring

Machine-learning models were retained as complementary predictive BI analysis rather than as substitutes for the PLS model. The target was the respondent-level arithmetic mean of CSV1–CSV8, termed the CSV indicator-mean score. This outcome is not identical to the Mode B CSV composite. Antecedent indicators were predictors, while CSV1–CSV8 were excluded from the predictor set to prevent leakage.

The dataset was randomly divided into 70% training and 30% testing subsets. Hyperparameter selection used k-fold cross-validation within the training data, and the test set was reserved for final evaluation. Decision Tree, Support Vector Machine, Random Forest, and AdaBoost models were compared using mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and R^2 (Table 9). These metrics quantify within-sample predictive performance under the reported split; they do not establish causality or external campaign effectiveness.

Table 9. Prediction of the CSV indicator-mean score.

ML Method	MSE	RMSE	MAE	R^2
Decision Tree	0.101	0.318	0.219	0.868
SVM	0.058	0.240	0.144	0.925
Random Forest	0.037	0.192	0.135	0.952
AdaBoost	0.036	0.190	0.122	0.953

AdaBoost achieved the lowest errors and highest test-set R^2 , with Random Forest producing nearly identical results. The high values are treated as internal predictive evidence from one moderate-sized sample and require external validation in an independent sample.

For an illustrative BI prioritization score, the five formative-model antecedent coefficients were normalized to sum approximately to one: $R_i = 0.261\text{Clarity}_i + 0.177\text{Likability}_i + 0.130\text{Empathy}_i + 0.205\text{Similarity}_i + 0.227\text{WTU}_i$. The score ranks anonymous profiles and illustrates how explanatory estimates may inform decision support; it is not a validated optimization rule, a formal MCDM model, or a causal estimate of conversion. Developing and externally validating an MCDM formulation is reserved for future research.

5. Discussion

This study examined how five content- and creator-related perceptions are associated with CSV and how CSV is associated with self-reported purchase intention among 409 Bulgarian respondents. The principal measurement correction treats CSV1–CSV8 as formative components: distinct product/economic, community/social, and relational contributions jointly define CSV [25–27].

The formative measurement evidence is acceptable but qualified. Six weights were significant, all absolute loadings exceeded 0.72, and all VIFs were below 5. CSV4 and CSV5 did not add significant unique weight after the other indicators were controlled, but their loadings and content coverage supported retention. Because no global CSV criterion was

included, redundancy validity could not be assessed. The former reflective reliability and AVE statistics therefore cannot be used as evidence for CSV; they remain relevant only to the reflective sensitivity model.

The structural results support all six prespecified associations. The five antecedents explained 69.1% of CSV variance, and CSV explained 32.2% of purchase-intention variance. Clarity was the largest antecedent ($\beta = 0.260$), followed by willingness to use ($\beta = 0.226$), similarity ($\beta = 0.205$), likability ($\beta = 0.177$), and empathy ($\beta = 0.130$). These estimates indicate that clear, understandable, and memorable communication and readiness to engage with SFV content are most strongly associated with attributed value.

CSV was positively associated with purchase intention ($\beta = 0.567$, $p < 0.001$), and Q^2 predict was positive for both retained purchase-intention indicators. This result is consistent with prior short-video social-commerce research connecting creator perceptions, shared value, and purchase intention [2]. However, the cross-sectional survey does not demonstrate that changing a content attribute will cause purchases. Competing models also showed direct associations between willingness to use and empathy and purchase intention. CSV remains positively associated with purchase intention within the hypothesized model, but the analyses do not independently establish mediation.

The findings refine the SOR interpretation by indicating that attributed value in SFV commerce encompasses both information-processing and relational evaluations. The comparatively stronger associations of clarity and willingness to use suggest that CSV is related not only to affective or relational responses to creators but also to whether their content is understandable and usable as an informational or decision-support resource. Relational perceptions such as empathy, similarity, and likability appear to provide additional, conceptually distinct contributions to this value appraisal. CSV therefore represents a broader organism-level construct than trust or attitude alone because it integrates product/economic, community/social, and relational contributions within a formative composite. In relation to TAM-based models, the framework distinguishes readiness to use SFV content from the multidimensional value attributed to the creator by treating them as conceptually separate constructs. Comparative models incorporating trust, attitude, and perceived value are nevertheless needed before concluding that CSV provides stronger explanatory performance than these alternative constructs. The cross-sectional findings support these conceptual distinctions but do not establish their causal or temporal ordering.

The relative ordering of the antecedents offers additional insight when considered alongside previous research on consumer responses to short-form video marketing. The results are particularly consistent with Shen and Wang [2], who connected clarity, likability, empathy, similarity, and willingness to use with shared-value creation and linked shared-value creation to purchase intention. In the present study, clarity has the strongest unique association with CSV. In rapidly changing short-video feeds, understandable and memorable content may reduce processing effort and uncertainty, making creators' product-related and social contributions easier to recognize. This interpretation may be particularly relevant to the comparatively young sample: 76.6% of respondents were aged 30 or younger, indicating substantial Generation Z representation, although the available age categories do not permit its precise proportion to be determined. In the context of this comparatively young and digitally experienced sample, consumers frequently encounter commercial information in rapidly changing, algorithmically curated feeds; consequently, immediate comprehensibility may be particularly important when deciding whether creator content merits further attention. This interpretation is consistent with previous research on Generation Z consumers demonstrating the relevance of interesting content, perceived usefulness, trust, and brand attitude in short-video marketing [6]. However, because age was not tested

as a moderator, the sample's age composition provides contextual support rather than evidence that the clarity–CSV association is stronger among Generation Z respondents.

Willingness to use has the second-largest association, suggesting that attributed value is related not only to favorable impressions but also to consumers' readiness to employ SFV content as an informational or decision-support resource. In the context of the comparatively young sample, usability may help consumers distinguish practically valuable creator content from content that is merely entertaining. Empathy may contribute by strengthening perceived understanding and relational connection with the creator, thereby supporting the relational component of CSV. Such relational cues may be particularly relevant in younger consumers' social-media environments, where creator interaction, perceived authenticity, and identification are prominent features of the platform experience. Nevertheless, empathy's smaller coefficient indicates a positive but more limited unique contribution after clarity, willingness to use, similarity, and likability are considered simultaneously. These age-related interpretations remain contextual because the study identified age-group differences in the outcome means but did not test whether age moderates the structural associations.

The group-difference analyses provide bounded support for H7. Across the 18 omnibus ANOVAs, both the CSV indicator mean and the purchase-intention mean differed significantly by age and followed-influencer count after Holm correction. Mean purchase intention also differed significantly by social-media-use frequency and daily SFV viewing time. Effect sizes for the significant findings ranged from $\eta^2 = 0.051$ to 0.097 . Thus, H7 is partially supported at the level of observed between-group mean differences. These results indicate heterogeneity in the outcome means but do not establish that demographic or usage characteristics caused the observed differences. Structural-path heterogeneity was not examined and is therefore not inferred.

The complementary BI analyses add context without changing the evidential status of the primary model. Clustering identifies broadly favorable and unfavorable response profiles but may partly capture general positivity. Sentiment analysis highlights clarity, practical information, authenticity, and disclosure, alongside distraction and distrust. The ML models predict the equal-weight CSV indicator mean within the reported split, but the target differs from the formative PLS score and has not been externally validated.

For BI-supported e-commerce management, the findings offer testable priorities rather than confirmed optimization rules. Message clarity and consumers' readiness to engage with and use SFV content merit particular attention. These suggestions should be tested with campaign experiments, clickstream records, and sales outcomes before operational deployment.

Several limitations constrain interpretation. The convenience sample was self-selected, concentrated in Bulgaria, and comparatively young; a response rate and nonresponse bias estimate were unavailable. All focal constructs were collected from the same respondents at one time, so common-method variance cannot be ruled out. The H7 ANOVAs retained original categories, and several infrequent-use or older-age cells were small; the social-media-frequency result therefore warrants particular caution. Purchase intention is not actual behavior, platform exposure was not observed, and no independent global CSV item was available for redundancy validation. Finally, the clustering and ML results require replication in independent samples.

The obtained results indicate that CSV can be represented as a perception-based formative composite and that its positive association with purchase intention remains stable across the reported sensitivity specifications.

6. Conclusions and Future Research

This study examined the associations among SFV perceptions, CSV, and purchase intention in a cross-sectional sample of 409 Bulgarian respondents. CSV was measured using eight indicators covering product/economic, community/social, and relational contributions associated with short-form video creators.

All five antecedents were positively associated with CSV, which was positively associated with purchase intention ($\beta = 0.567$). The model explained 69.1% of the variance in CSV and 32.2% of the variance in purchase intention, while the Q^2 predict values for PI1 and PI3 were positive. The measurement-model assessment supported the inclusion of the selected indicators in the reported model. Overall, the findings support the proposed associations, although the cross-sectional design does not permit causal conclusions.

The Holm-adjusted ANOVAs partially supported H7: age and followed-influencer count distinguished both outcome means, whereas social-media use frequency and daily viewing time distinguished purchase intention. These findings concern differences in observed outcome means; structural-path heterogeneity was not examined.

The paper's contribution is empirical, measurement-oriented, and BI-relevant rather than an implementation-ready workflow. It shows how descriptive profiling, formative structural modeling, segmentation, sentiment analysis, group-difference testing, predictive machine learning, and transparent scoring can convert consumer-response data into complementary e-commerce insights without treating statistical associations or internal predictions as causal evidence of campaign effectiveness.

Future studies should include an independent global CSV item for redundancy validity, compare one- and multidimensional formative specifications, and preregister direct and mediated paths. Probability-based or more diverse samples, longitudinal designs, field experiments, and behavioral measures such as watch time, clicks, saves, purchases, and repeat purchases are needed to assess causality, external validity, and actual commercial outcomes. Future work should also develop and validate a formal MCDM model that combines explanatory, predictive, and managerial criteria rather than treating the present illustrative score as a completed MCDM application.

Author Contributions: Conceptualization, G.I., T.Y., M.R., D.A., S.K.-B., M.B., P.G. and A.D.; modeling, G.I., T.Y., M.R. and S.K.-B.; validation, G.I., T.Y., M.R. and M.B.; formal analysis, T.Y.; resources, G.I., T.Y., M.R. and S.K.-B.; writing—original draft preparation, G.I.; writing—review and editing, G.I., T.Y. and P.G.; visualization, T.Y. and S.K.-B.; supervision, G.I.; project administration, M.R. and A.D.; funding acquisition, G.I., T.Y. and S.K.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Project BG16RFPR002-1.014-0013-C01 “Digitalization of Economy in Big Data Environment– Second Stage” (DIGD2) financed by the “Research, Innovation and Digitalization for Smart Transformation” Program 2021–2027 and co-funded by the European Union.

Institutional Review Board Statement: Ethical review and approval were not required for this study, as it did not involve any sensitive or personally identifiable data. Participation was voluntary and anonymous. All procedures complied with the EU General Data Protection Regulation (GDPR) and applicable national legislation.

Informed Consent Statement: An informed consent statement was not required because the study posed no risk to participants, and data were collected anonymously without any personal identifiers, in accordance with GDPR and relevant national regulations.

Data Availability Statement: The data stored as csv and pdf files are publicly available at <https://data.mendeley.com/datasets/vhzzfj8hdf/1> (accessed on 1 July 2026).

Acknowledgments: The authors thank the academic editor and anonymous reviewers for their insightful comments and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Luo, C.; Mohd Hasan, N.A.; Zamri bin Ahmad, A.M.; Lei, G. Influence of short video content on consumers purchase intentions on social media platforms with trust as a mediator. *Sci. Rep.* **2025**, *15*, 16605. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Shen, X.; Wang, J. How short video marketing influences purchase intention in social commerce: The role of users' persona perception, shared values, and individual-level factors. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 290. [\[CrossRef\]](#)
3. Dwinanda, B.; Syaripuddin, F.A.; Hudaifi; Hendriana, E. Examining the extended advertising value model: A case of TikTok short video ads. *Mediterr. J. Soc. Behav. Res.* **2022**, *6*, 35–44. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Zhai, L.; Yin, P.; Li, C.; Wang, J.; Yang, M. Investigating the effects of video-based e-word-of-mouth on consumers' purchase intention: The moderating role of involvement. *Sustainability* **2022**, *14*, 9522. [\[CrossRef\]](#)
5. Yin, X.; Li, Y.; Gao, R.; Li, J.; Wang, H. Understanding the purchase decisions of silver consumers in short-form video platforms from the perspective of existence, relatedness, and growth needs. *Behav. Sci.* **2023**, *13*, 1011. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Ngo, T.T.A.; Quach, P.; Nguyen, T.V.; Nguyen, A.D.; Nguyen, T.M.N. Short video marketing factors influencing the purchase intention of Generation Z in Vietnam. *Innov. Mark.* **2023**, *19*, 34–50. [\[CrossRef\]](#)
7. Porter, M.E.; Kramer, M.R. Creating shared value. *Harv. Bus. Rev.* **2011**, *89*, 62–77.
8. Ajzen, I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **1991**, *50*, 179–211. [\[CrossRef\]](#)
9. Morwitz, V. Consumers' purchase intentions and their behavior. *Found. Trends Mark.* **2014**, *7*, 181–230. [\[CrossRef\]](#)
10. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*, 319–340. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User acceptance of information technology: Toward a unified view. *MIS Q.* **2003**, *27*, 425–478. [\[CrossRef\]](#)
12. Shmueli, G.; Koppius, O.R. Predictive analytics in information systems research. *MIS Q.* **2011**, *35*, 553–572. [\[CrossRef\]](#)
13. Hair, J.F.; Hult, G.T.M.; Ringle, C.M.; Sarstedt, M.; Danks, N.P.; Ray, S. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 3rd ed.; Sage: Thousand Oaks, CA, USA, 2021.
14. Yin, X.; Li, J.; Si, H.; Wu, P. Attention marketing in fragmented entertainment: How advertising embedding influences purchase decision in short-form video apps. *J. Retail. Consum. Serv.* **2024**, *76*, 103572. [\[CrossRef\]](#)
15. Meng, L.; Kou, S.; Duan, S.; Bie, Y. The impact of content characteristics of short-form video ads on consumer purchase behavior: Evidence from TikTok. *J. Bus. Res.* **2024**, *183*, 114874. [\[CrossRef\]](#)
16. Feng, Y.; Zhai, M.; Du, Y. Impact mechanisms of mini-detail short videos on consumer purchase intention on Taobao: A TAM2-based approach. *Entertain. Comput.* **2024**, *51*, 100745. [\[CrossRef\]](#)
17. Jiang, W.; Chen, H.-L. Can short videos work? The effects of use and gratification and social presence on purchase intention: Examining the mediating role of digital dependency. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 5. [\[CrossRef\]](#)
18. Yu, S.; Zhang, H.; Zheng, Q.; Chu, D.D.; Chen, T.; Chen, X. Consumer behavior based on the SOR model: How do short video advertisements affect furniture consumers' purchase intentions? *BioResources* **2024**, *19*, 2639–2659. [\[CrossRef\]](#)
19. Mehrabian, A.; Russell, J.A. *An Approach to Environmental Psychology*; MIT Press: Cambridge, MA, USA, 1974.
20. Petty, R.E.; Cacioppo, J.T. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*; Springer: New York, NY, USA, 1986.
21. Venkatesh, V.; Davis, F.D. A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Manag. Sci.* **2000**, *46*, 186–204. [\[CrossRef\]](#)
22. Katz, E.; Blumler, J.G.; Gurevitch, M. Uses and gratifications research. *Public Opin. Q.* **1973**, *37*, 509–523. [\[CrossRef\]](#)
23. Short, J.; Williams, E.; Christie, B. *The Social Psychology of Telecommunications*; Wiley: London, UK, 1976.
24. Yu, S.; Wu, Z. Research on the influence mechanism of short video communication effect of furniture brand: Based on ELM model and regression analysis. *BioResources* **2024**, *19*, 3191–3207. [\[CrossRef\]](#)
25. Jarvis, C.B.; MacKenzie, S.B.; Podsakoff, P.M. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *J. Consum. Res.* **2003**, *30*, 199–218. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Diamantopoulos, A.; Winklhofer, H.M. Index construction with formative indicators: An alternative to scale development. *J. Mark. Res.* **2001**, *38*, 269–277. [\[CrossRef\]](#)
27. Hair, J.F.; Hult, G.T.M.; Ringle, C.M.; Sarstedt, M.; Danks, N.P.; Ray, S. Evaluation of formative measurement models. In *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*; Springer: Cham, Switzerland, 2021. [\[CrossRef\]](#)
28. Yoon, S.-H.; Lee, S.-H. What likeability attributes attract people to watch online video advertisements? *Electronics* **2022**, *11*, 1960. [\[CrossRef\]](#)

29. Cuff, B.M.P.; Brown, S.J.; Taylor, L.; Howat, D.J. Empathy: A review of the concept. *Emot. Rev.* **2016**, *8*, 144–153. [CrossRef]
30. Decety, J.; Jackson, P.L. The functional architecture of human empathy. *Behav. Cogn. Neurosci. Rev.* **2004**, *3*, 71–100. [CrossRef] [PubMed]
31. Li, B.; Chen, S.; Zhou, Q. Empathy with influencers? The impact of the sensory advertising experience on user behavioral responses. *J. Retail. Consum. Serv.* **2023**, *72*, 103286. [CrossRef]
32. Mangus, S.M.; Bock, D.E.; Jones, E.; Folse, J.A.G. Examining the effects of mutual information sharing and relationship empathy: A social penetration theory perspective. *J. Bus. Res.* **2020**, *109*, 375–384. [CrossRef]
33. Yuan, S.; Lou, C. How social media influencers foster relationships with followers: The roles of source credibility and fairness in parasocial relationship and product interest. *J. Interact. Advert.* **2020**, *20*, 133–147. [CrossRef]
34. Aparicio, M.; Costa, C.; Moises, R. Gamification and reputation: Key determinants of e-commerce usage and repurchase intention. *Heliyon* **2021**, *7*, e06383. [CrossRef] [PubMed]
35. Ham, S.; Lee, S.; Yoon, H.J.; Kim, C.S. Linking creating shared value to customer behaviors in the food service context. *J. Hosp. Tour. Manag.* **2020**, *43*, 199–208. [CrossRef]
36. Dillman, D.A.; Smyth, J.D.; Christian, L.M. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, 4th ed.; Wiley: Hoboken, NJ, USA, 2014.
37. DeVellis, R.F.; Thorpe, C.T. *Scale Development: Theory and Applications*, 5th ed.; Sage: Thousand Oaks, CA, USA, 2022.
38. Likert, R. A technique for the measurement of attitudes. *Arch. Psychol.* **1932**, *22*, 55.
39. Chu, S.K.W.; Huang, H.; Wong, W.; Van Ginneken, W.N.M.; Wu, K.M.; Hung, M.Y. Quality and clarity of health information on Q&A sites. *Libr. Inf. Sci. Res.* **2018**, *40*, 237–244. [CrossRef]
40. Bornet, C.; Brangier, E. The effects of personas on creative codesign of work equipment: An exploratory study in a real setting. *CoDesign* **2016**, *12*, 243–256. [CrossRef]
41. Xiao, C.C.; Zhu, Y.R.; Hu, D.; Xie, Z.Q. The impact of TikTok video marketing on tourist destination image cognition and tourism intention. In Proceedings of the 2020 International Workshop on Electronic Communication and Artificial Intelligence, Shanghai, China, 12–14 June 2020; pp. 116–119. [CrossRef]
42. Kim, H.-W.; Chan, H.C.; Kankanalli, A. What motivates people to purchase digital items on virtual community websites? The desire for online self-presentation. *Inf. Syst. Res.* **2012**, *23*, 1232–1245. [CrossRef]
43. Field, A. *Discovering Statistics Using IBM SPSS Statistics*, 5th ed.; Sage: London, UK, 2018.
44. Ringle, C.M.; Wende, S.; Becker, J.-M. *SmartPLS 3*; SmartPLS GmbH: Bönningstedt, Germany, 2015.
45. Dijkstra, T.K.; Henseler, J. Consistent partial least squares path modeling. *MIS Q.* **2015**, *39*, 297–316. [CrossRef]
46. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
47. Hayes, A.F. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, 2nd ed.; Guilford Press: New York, NY, USA, 2018.
48. Becker, J.-M.; Ringle, C.M.; Sarstedt, M. Estimating moderating effects in PLS-SEM and PLSc-SEM: Interaction term generation and data treatment. *J. Appl. Struct. Equ. Model.* **2018**, *2*, 1–21. [CrossRef] [PubMed]
49. Sarstedt, M.; Henseler, J.; Ringle, C.M. Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. *Adv. Int. Mark.* **2011**, *22*, 195–218. [CrossRef]
50. Henseler, J.; Ringle, C.M.; Sarstedt, M. Testing measurement invariance of composites using partial least squares. *Int. Mark. Rev.* **2016**, *33*, 405–431. [CrossRef]
51. MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1967; Volume 1, pp. 281–297.
52. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [CrossRef]
53. Tzeng, G.-H.; Huang, J.-J. *Multiple Attribute Decision Making: Methods and Applications*; CRC Press: Boca Raton, FL, USA, 2011.
54. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*, 2nd ed.; Springer: New York, NY, USA, 2021. [CrossRef]
55. Pang, B.; Lee, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 1–135. [CrossRef]
56. Ilieva, G.; Yankova, T.; Ruseva, M.; Angelov, D.; Klisarova-Belcheva, S.; Bratkov, M.; Dimitrov, A. Consumer Attitudes towards Short Videos in Social Media Platforms. Mendeley Data, V1, 2026. Available online: <https://data.mendeley.com/datasets/vhzzfj8hdf/1> (accessed on 1 July 2026).
57. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2024. Available online: <https://www.R-project.org/> (accessed on 30 June 2026).
58. Kassambara, A.; Mundt, F. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses, R Package Version 1.0.7, 2020. Available online: <https://CRAN.R-project.org/package=factoextra> (accessed on 30 June 2026).
59. DataReportal. Digital 2025: Bulgaria. Available online: <https://datareportal.com/reports/digital-2025-bulgaria> (accessed on 30 June 2026).

60. Demšar, J.; Curk, T.; Erjavec, A.; Gorup, Č.; Hočevár, T.; Milutinovič, M.; Možina, M.; Polajnar, M.; Toplak, M.; Starič, A.; et al. Orange: Data mining toolbox in Python. *J. Mach. Learn. Res.* **2013**, *14*, 2349–2353.
61. Loria, S. TextBlob: Simplified Text Processing, 2018. Available online: <https://textblob.readthedocs.io/> (accessed on 30 June 2026).
62. Henseler, J.; Ringle, C.M.; Sarstedt, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* **2015**, *43*, 115–135. [[CrossRef](#)]
63. Hair, J.F.; Risher, J.J.; Sarstedt, M.; Ringle, C.M. When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev.* **2019**, *31*, 2–24. [[CrossRef](#)]
64. Fornell, C.; Larcker, D.F. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* **1981**, *18*, 39–50. [[CrossRef](#)]
65. Stone, M. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B* **1974**, *36*, 111–133. [[CrossRef](#)]
66. Geisser, S. A predictive approach to the random effects model. *Biometrika* **1974**, *61*, 101–107. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.