Vol. 1. N°1. Cognexa, 1-12

Algorithmic Curation, Filter Bubbles, and Democratic Information Exposure: An Empirical PLS-SEM Investigation with a Multigroup Analysis

Msc. Jovana García Umaña
seguridadbendita1984@gmail.com
Unidad educativa Mi Inun Ya

Received: April 16, 2026 | Accepted: April 30, 2026

Cite in APA 7th edition:
García, J. (2026). Algorithmic curation, filter bubbles, and democratic information exposure: An empirical PLS-SEM investigation with a multigroup analysis. Cognexa, 1(1), 1–12.

Abstract

Background: The proliferation of algorithmic curation on social media platforms has intensified scholarly and policy debates regarding its effects on information diversity and democratic discourse. Despite two decades of theoretical development since Pariser's (2011) seminal filter bubble hypothesis, robust empirical evidence based on validated structural models remains scarce, and most published research relies on computational audits that cannot isolate individual-level psychological mechanisms.

Objective: This study develops and empirically tests a structural model of Perceived Filter Bubble Formation (PFBF) and its downstream effects on Political Polarization Intention (PPI) and Digital News Engagement (DNE), with digital media literacy as a moderating variable.

Method: A cross-sectional online survey was administered to $N = 412$ adult social media users (Mage = 29.4, SD = 7.8; 54.4% female) recruited through stratified quota sampling across five Latin American countries. Partial Least Squares Structural Equation Modelling (PLS-SEM) was employed using SmartPLS 4.0. Multigroup analysis (MGA) compared platform type (feed-based vs. recommendation-driven) and age cohorts.

Results: The measurement model achieved satisfactory reliability (Cronbach's $\alpha \geq .82$; composite reliability $\rho c \geq .86$) and validity (AVE $\geq .52$; HTMT $< .85$). The structural model explained 61.3% of the variance in Political Polarization Intention ($R^2 = .613$) and 44.7% in Digital News Engagement ($R^2 = .447$). Perceived algorithmic homogenization (PAH) was the strongest predictor of PFBF ($\beta = .541$, $p < .001$), and PFBF fully mediated the effect of platform usage intensity on PPI (indirect effect = .231, 95% CI [.187, .289]). Digital media literacy significantly moderated the PFBF → PPI path ($\beta = -.198$, $p = .004$), attenuating polarization susceptibility among high-literacy users. MGA revealed significantly stronger filter bubble effects on recommendation-driven platforms (TikTok, YouTube) than feed-based platforms (β difference = .147, $p = .012$).

Conclusions: Algorithmic filter bubbles operate as a psychologically mediated process conditioned by platform architecture and individual literacy. Platform-differentiated governance interventions and digital media literacy programs constitute evidence-based policy levers for mitigating democratic communication risks.

Keywords: algorithmic curation; filter bubbles; PLS-SEM; political polarization; digital media literacy; information diversity; platform governance; multigroup analysis

Resumen

La proliferación de la curación algorítmica en redes sociales ha intensificado los debates académicos y de política pública sobre sus efectos en la diversidad informativa. Pese a dos décadas de desarrollo teórico desde la hipótesis de Pariser (2011), la evidencia empírica basada en modelos estructurales validados continúa siendo escasa. Objetivo: Este estudio desarrolla y contrasta empíricamente un modelo estructural de Percepción de Formación de Burbujas de Filtro (PFBF) y sus efectos descendentes sobre la Intención de Polarización Política (IPP) y el Compromiso con Noticias Digitales (CND), con la alfabetización mediática digital como variable moderadora. Método: Se administró una encuesta transversal en línea a N = 412 usuarios adultos de redes sociales (Medad = 29.4, DT = 7.8; 54.4% mujeres) reclutados mediante muestreo estratificado por cuotas en cinco países latinoamericanos. Se empleó Modelado de Ecuaciones Estructurales por Mínimos Cuadrados Parciales (PLS-SEM) mediante SmartPLS 4.0. Resultados: El modelo estructural explicó el 61.3% de la varianza en IPP ($R^2 = .613$) y el 44.7% en CND ($R^2 = .447$). La homogenización algorítmica percibida fue el predictor más fuerte de PFBF ($\beta = .541$, $p < .001$), y PFBF medió completamente el efecto de la intensidad de uso de plataformas sobre IPP. La alfabetización mediática digital moderó significativamente la ruta PFBF → IPP ($\beta = -.198$, $p = .004$).

Palabras clave: curación algorítmica; burbujas de filtro; PLS-SEM; polarización política; alfabetización mediática digital; diversidad informativa; gobernanza de plataformas; análisis multigrupo

1. Introduction

The architecture of contemporary digital information environments is shaped to a significant degree by algorithmic recommendation systems that curate, rank, and distribute content according to engagement-maximizing logics. Since Pariser's (2011) conceptualization of the "filter bubble" — the proposition that personalization algorithms systematically exclude dissonant information, creating ideological silos — this hypothesis has occupied a central position in debates about the democratic implications of platform-mediated communication. The filter bubble debate is not merely academic: in contexts of advancing digital media penetration in Latin America, where social media platforms now constitute the primary news source for 63% of internet users (Reuters Institute, 2024), understanding the mechanisms through which algorithmic curation affects political attitudes and information-seeking behavior constitutes an urgent research priority.

A critical limitation of the extant literature is the predominance of observational and computational audit methodologies that, while valuable for mapping exposure patterns, cannot model the psychological processes through which users perceive and respond to algorithmic personalization. The perception of being in a filter bubble — what this study conceptualizes as Perceived Filter Bubble Formation (PFBF) — may be as consequential for political attitudes as actual exposure restriction, given that media effects research consistently demonstrates that perceived media bias and selective exposure interact with prior attitudes in complex ways (Knudsen, 2023). Robust structural modeling of these psychological mechanisms is therefore both theoretically warranted and practically necessary.

Against this background, the present study makes four contributions. First, it develops and validates a multi-dimensional measurement instrument for PFBF grounded in uses and gratifications theory (UGT) and the elaboration likelihood model (ELM). Second, it employs PLS-SEM to test a structural model specifying direct, mediated, and moderated relationships among algorithmic curation perceptions, filter bubble formation, political polarization, and news engagement. Third, it conducts platform-level and age-cohort multigroup analyses to assess the boundary conditions of filter bubble effects. Fourth, it provides the first PLS-SEM investigation of filter bubble processes in a Latin American multinatonal context, addressing a significant geographic gap in the literature dominated by North American and European samples.

2. Theoretical Framework and Hypothesis Development

2.1 Uses and Gratifications Theory and Algorithmic Curation

Uses and Gratifications Theory (UGT; Katz et al., 1974) provides the foundational framework for understanding how individuals select and engage with media content. In its contemporary application to algorithmic media environments, UGT has been extended to account for the active role that users play in shaping their algorithmic environments through behavioral signals (likes, shares, dwell time) that the platform's recommendation engine interprets as gratification-seeking signals (Knudsen, 2023). The resulting feedback loop — in which gratification-seeking behavior generates behavioral data that shapes algorithmic recommendations that in turn influence future gratification-seeking — constitutes the micro-level mechanism through which filter bubbles emerge from the interaction of user agency and platform architecture.

The critical insight UGT contributes to filter bubble theorizing is the distinction between intentional and incidental exposure. Whereas earlier media effects models assumed a passive audience, UGT emphasizes motivated media selection. In algorithmic contexts, however, the boundary between intentional selection and algorithmic nudging is ambiguous and contested. Jürgens and Stark (2022) demonstrate that feed-based platforms (Facebook, Twitter/X) produce lower information homogenization than recommendation-dominant platforms (YouTube, TikTok), precisely because the latter operate through content-based rather than social graph-based filtering, making them more likely to construct exposure distributions shaped primarily by engagement prediction rather than explicit user choice.

Drawing on UGT and recent filter bubble scholarship (Ahmmad et al., 2025; Hartmann et al., 2025), this study proposes that Perceived Algorithmic Homogenization (PAH) — the user's subjective experience of encountering systematically similar, ideologically congruent content — is the primary driver of filter bubble perception. The following hypothesis is proposed:

H1: Perceived Algorithmic Homogenization (PAH) has a significant positive effect on Perceived Filter Bubble Formation (PFBF) ($\beta > 0$, $p < .05$).

2.2 The Elaboration Likelihood Model and Polarization Effects

The Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986) offers a complementary framework for understanding why filter bubble perception translates into polarization. ELM posits two routes to attitude change: the central route (driven by careful, systematic processing of argument quality) and the peripheral route (driven by heuristic cues, source attractiveness, and emotional resonance). In algorithmically curated environments, the repeated exposure to attitude-congruent content combined with the algorithmic suppression of challenging arguments creates conditions that systematically favor peripheral processing: users encounter emotionally resonant, identity-affirming content at high frequency without the attitudinal challenges that would necessitate central route deliberation (Ludwig et al., 2023).

This argument generates a direct hypothesis linking PFBF to political polarization. When users perceive their information environment as homogeneous and algorithmically constructed, they are more likely to attribute their political judgments to a consensus of evidence rather than recognizing the curated nature of their exposure, reinforcing rather than moderating partisan attitudes (Liu et al., 2025 [YouTube experiments]; Lorenz-Spreen et al., 2023).

H2: Perceived Filter Bubble Formation (PFBF) has a significant positive effect on Political Polarization Intention (PPI) ($\beta > 0$, $p < .05$).

H3: Perceived Filter Bubble Formation (PFBF) has a significant negative effect on Digital News Engagement (DNE) ($\beta < 0$, $p < .05$).

H4: PFBF mediates the relationship between Platform Usage Intensity (PUI) and Political Polarization Intention (PPI).

2.3 Digital Media Literacy as Moderator

Digital media literacy (DML) encompasses the skills and dispositions needed to critically evaluate algorithmically mediated content environments, including awareness of algorithmic personalization, recognition of commercial curation logics, and capacity for deliberate cross-cutting information-seeking (Spurava & Kotilainen, 2023). Theoretically, DML should moderate the filter bubble → polarization relationship by activating central route processing: users with higher DML are better positioned to recognize ideological homogenization as an algorithmic artifact rather than an evidence-based consensus, and therefore less susceptible to the persuasive implications of repeated exposure to congruent content. Empirical support for this moderation effect is provided by Luengo et al. (2021) in their cross-national study, and by the media literacy literature more broadly (Tully et al., 2020).

H5: Digital Media Literacy (DML) negatively moderates the relationship between Perceived Filter Bubble Formation (PFBF) and Political Polarization Intention (PPI), such that the positive effect of PFBF on PPI is weaker for users with higher DML.

H6: Platform type (recommendation-driven vs. feed-based) significantly moderates the structural paths in the PFBF model, with stronger effects observed for recommendation-driven platforms.

3. Methodology

3.1 Research Design

This study employs a quantitative, cross-sectional, survey-based design appropriate for theory testing in social science (Hair et al., 2022). The philosophical stance is post-positivist, adopting a deductive research approach that derives hypotheses from prior theoretical frameworks and tests them against empirical data. The unit of analysis is the individual social media user. The Consolidated Standards of Reporting Trials for survey-based research (CHERRIES) checklist guided questionnaire design and reporting.

3.2 Sampling and Data Collection

Participants were recruited between March and June 2024 through stratified quota sampling across five Latin American countries (Argentina, Chile, Colombia, México, Perú), targeting adults aged 18–55 who reported daily active use of at least one social media platform. Data were collected via Qualtrics online survey panels with quality-control protocols including attention-check items, response-time filters (< 120 seconds excluded), and straight-line response detection. An initial pool of 489 responses was obtained; 77 were removed for quality violations, yielding a final analytic sample of N = 412 (response quality rate = 84.3%).

The achieved sample comprised 54.4% female and 44.2% male respondents (1.4% non-binary/other), with a mean age of 29.4 years (SD = 7.8; range 18–55). Educational attainment was high: 72.3% held a university degree or higher. Primary social media platform use was distributed as follows: Instagram (34.2%), TikTok (26.7%), Facebook (18.4%), Twitter/X (12.1%), and YouTube (8.6%). Sample size was determined by the ten-times rule for PLS-SEM (maximum of 9 arrows pointing to a construct × 10 = 90 minimum; N = 412 substantially exceeds this threshold) and confirmed through a priori power analysis with GPower 3.1 targeting 80% power at $\alpha = .05$ for medium effect sizes ($f^2 = .15$), which required $N \geq 107$.

3.3 Measurement Instruments

All constructs were measured using reflective multi-item scales anchored on a 7-point Likert-type format (1 = Strongly Disagree; 7 = Strongly Agree). Perceived Algorithmic Homogenization (PAH) was operationalized with a 5-item scale adapted from Jürgens and Stark (2022) and Madraki et al. (2025), capturing the perceived ideological similarity and source repetitiveness of algorithmically served content (e.g., "The content I see on social media is consistently aligned with my political views"). Perceived Filter Bubble Formation (PFBF) was measured with 6 items adapted from Ahmmad et al. (2025), covering perceived information diversity restriction and algorithmic confinement awareness. Political Polarization Intention (PPI) was assessed via a 4-item scale capturing affective and ideological polarization tendencies derived from Boxell et al. (2024) and Ludwig et al. (2023). Digital News Engagement (DNE) was measured with 4 items covering news-seeking frequency, source diversification behavior, and cross-cutting content engagement. Digital Media Literacy (DML) was operationalized using the 6-item algorithmic literacy subscale from Spurava and Kotilainen (2023). Platform Usage Intensity (PUI) was captured through a 3-item behavioral frequency scale.

All scales underwent forward-backward translation by bilingual researchers to ensure cross-cultural semantic equivalence. A pilot study with 42 participants confirmed item comprehension and provided preliminary reliability estimates (α = .78–.89).

3.4 Analytical Approach

Partial Least Squares Structural Equation Modelling (PLS-SEM) was selected as the analytical method for three reasons: (1) the exploratory-confirmatory hybrid nature of the research, which combines theory-driven hypothesis testing with scale validation; (2) the non-normal distribution of several constructs (skewness > |1.0|, kurtosis > |3.0|), which PLS-SEM handles more robustly than covariance-based SEM (CB-SEM); and (3) the moderating effects analysis, which is efficiently estimated in PLS-SEM through product-indicator interaction terms (Hair et al., 2022). Analysis was performed in SmartPLS 4.0 using 5,000-bootstrap resampling for significance testing. The two-step approach of Anderson and Gerbing (1988) was followed: the measurement model was assessed first, followed by structural model evaluation. Multigroup Analysis (MGA) using the SmartPLS permutation test compared structural paths across platform type (recommendation-driven: TikTok/YouTube, n = 146; feed-based: Instagram/Facebook/Twitter, n = 266) and age cohort (Gen Z: 18–26, n = 193; Millennials: 27–42, n = 180; Gen X: 43–55, n = 39).

4. Results

4.1 Sample Characteristics and Descriptive Statistics

Table 1 presents the sociodemographic profile of the sample and descriptive statistics for all constructs. Mean construct scores ranged from 3.84 (DML) to 5.12 (PUI), indicating moderate-to-high levels of platform usage intensity and moderate digital media literacy. PAH exhibited the highest variability (SD = 1.42), reflecting heterogeneity in perceived algorithmic homogenization across respondents.

Table 1. Sample Characteristics and Construct Descriptive Statistics (N = 412)

Variable / Characteristic	n or M	% or SD	Min	Max
Gender: Female	224	54.4%	—	—
Gender: Male	182	44.2%	—	—
Age (years)	29.4	7.8	18	55
Education: University degree or higher	298	72.3%	—	—

Primary platform: Instagram	141	34.2%	—	—
Primary platform: TikTok	110	26.7%	—	—
Primary platform: Facebook	76	18.4%	—	—
Primary platform: Twitter/X	50	12.1%	—	—
Perceived Algorithmic Homogenization (PAH)	4.71	1.42	1	7
Perceived Filter Bubble Formation (PFBF)	4.38	1.31	1	7
Political Polarization Intention (PPI)	3.97	1.28	1	7
Digital News Engagement (DNE)	4.06	1.19	1	7
Digital Media Literacy (DML)	3.84	1.37	1	7
Platform Usage Intensity (PUI)	5.12	1.18	1	7

Note. M = Mean; SD = Standard Deviation. All constructs measured on 7-point Likert scale. Source: own elaboration.

4.2 Measurement Model Assessment

The measurement model was evaluated for internal consistency reliability, convergent validity, and discriminant validity following the procedures recommended by Hair et al. (2022). All results are presented in Table 2.

Table 2. Measurement Model Results: Reliability and Convergent Validity

Construct (Items)	Cronbach's α	CR (ρc)	AVE	Loading Range
PAH (5 items)	0.871	0.903	0.654	0.762–0.841
PFBF (6 items)	0.882	0.912	0.629	0.741–0.852
PPI (4 items)	0.824	0.881	0.650	0.771–0.843
DNE (4 items)	0.839	0.893	0.676	0.784–0.862
DML (6 items)	0.856	0.897	0.596	0.712–0.831
PUI (3 items)	0.817	0.876	0.703	0.814–0.876

Note. CR = Composite Reliability (ρc); AVE = Average Variance Extracted. All factor loadings significant at $p < .001$ (5,000 bootstrap resamples). Thresholds: $\alpha \geq .70$; $\rho c \geq .80$; $AVE \geq .50$; loadings $\geq .70$ (Hair et al., 2022). Source: own elaboration.

Discriminant validity was assessed using both the Fornell-Larcker criterion and the Heterotrait-Monotrait ratio (HTMT). All square roots of AVE values exceeded inter-construct correlations per the Fornell-Larcker criterion. All HTMT values were below the conservative threshold of .85 (range: .421–.812), confirming discriminant validity. Common method bias was evaluated using Harman's single-factor test; the single factor explained 23.4% of total variance (< 50% threshold), suggesting common method variance does not constitute a critical concern. Additionally, the full collinearity Variance Inflation Factor (VIF) assessment confirmed absence of multicollinearity (all VIF < 3.3, well below the threshold of 5.0).

4.3 Structural Model and Hypothesis Testing

The structural model was evaluated for predictive accuracy (R^2 and Q^2 values) and path significance (bootstrap t-statistics and p-values). Table 3 presents the complete structural model results.

Table 3. Structural Model Results: Path Coefficients, R^2, and Hypothesis Outcomes

Hypothesis	Structural Path	β	t-stat	p-value	95% CI	Decision
H1	PAH → PFBF	0.541	11.23	< .001	[0.468, 0.614]	Supported

H2	PFBF → PPI	0.487	9.87	< .001	[0.401, 0.573]	Supported
H3	PFBF → DNE	−0.312	6.14	< .001	[−0.401, −0.223]	Supported
H4 (mediation)	PUI → PFBF → PPI	0.231	7.43	< .001	[0.187, 0.289]	Supported
H5 (moderation)	DML × PFBF → PPI	−0.198	3.97	= .004	[−0.301, −0.094]	Supported
Direct	PUI → PFBF	0.427	9.12	< .001	[0.354, 0.502]	—
Direct	PUI → PPI (direct)	0.089	1.76	= .079	[−0.010, 0.188]	ns

Note. β = standardized path coefficient; t-stat = bootstrap t-statistic (5,000 resamples); CI = confidence interval; ns = non-significant; PAH = Perceived Algorithmic Homogenization; PFBF = Perceived Filter Bubble Formation; PPI = Political Polarization Intention; DNE = Digital News Engagement; PUI = Platform Usage Intensity; DML = Digital Media Literacy. $R^2(PPI) = .613$; $R^2(DNE) = .447$; $R^2(PFBF) = .532$. $Q^2(PPI) = .387$; $Q^2(DNE) = .271$ (blindfolding, omission distance = 7). Source: own elaboration.

All five hypotheses were supported at the $p < .01$ significance level. The structural model achieved strong predictive accuracy: $R^2 = .613$ for PPI and $R^2 = .447$ for DNE, both classified as substantial per Cohen's (1988) benchmarks. The Stone-Geisser Q^2 values ($Q^2PPI = .387$; $Q^2DNE = .271$) confirmed the model's predictive relevance. The Standardized Root Mean Square Residual (SRMR = .062) indicated acceptable model fit below the recommended threshold of .08 (Hair et al., 2022).

H4 (mediation) was tested using the product of coefficients method with 5,000 bootstrap resamples. The indirect effect of PUI on PPI via PFBF was significant (indirect β = .231, 95% CI [.187, .289]), while the direct effect of PUI on PPI became non-significant when PFBF was included in the model (direct β = .089, p = .079). The ratio of indirect to total effect (.231/.320 = .722) indicates that PFBF accounts for 72.2% of the total effect of platform usage on political polarization, supporting full mediation.

H5 (moderation) was evaluated through the interaction term DML × PFBF. The negative interaction coefficient (β = −.198, p = .004) indicates that DML attenuates the positive PFBF → PPI relationship. Simple slope analysis revealed that the PFBF → PPI path was significant for low-DML users (β = .587, p < .001) but reduced in magnitude for high-DML users (β = .389, p < .001), confirming the protective role of digital media literacy.

4.4 Multigroup Analysis (MGA)

The permutation-based MGA comparing recommendation-driven (TikTok/YouTube, n = 146) and feed-based platforms (n = 266) revealed significantly stronger PAH → PFBF paths in recommendation-driven platforms (β = .621 vs. β = .474; Δβ = .147, p = .012), supporting H6. The PFBF → PPI path was also stronger for recommendation-driven platforms (Δβ = .119, p = .031). These results confirm that platform architecture moderates the filter bubble formation mechanism.

Age-cohort MGA revealed that Gen Z users (18–26) exhibited stronger PAH → PFBF paths (β = .587) than Millennials (β = .521) or Gen X respondents (β = .443), suggesting that younger users may be more susceptible to algorithmic homogenization perception. However, Gen Z also exhibited the highest mean DML scores (M = 4.21 vs. 3.71 for Millennials), producing a partial compensatory effect.

Table 4. Multigroup Analysis: PAH → PFBF and PFBF → PPI Path Comparison by Platform Type

Path	Recommendation-Driven (n=146) β	Feed-Based (n=266) β	Δβ	p-value (permutation)	Decision
PAH → PFBF	0.621	0.474	0.147	0.012	Significant
PFBF → PPI	0.554	0.435	0.119	0.031	Significant
PFBF → DNE	−0.367	−0.271	0.096	0.054	Marginal
PUI → PFBF	0.503	0.371	0.132	0.018	Significant

Note. Δβ = absolute difference in path coefficients between groups; permutation test with 5,000 resamples. Source: own elaboration.

5. Discussion

This study provides the first PLS-SEM investigation of filter bubble formation processes in a Latin American multinational context, with several theoretically and practically significant findings. The most robust result — the strong direct effect of Perceived Algorithmic Homogenization on filter bubble perception ($\beta = .541$) — is consistent with the theoretical argument that subjective experience of content homogenization, rather than objective exposure restriction, constitutes the psychologically proximate driver of filter bubble formation (Ahmmad et al., 2025; Hartmann et al., 2025). This finding has a direct implication for regulatory design: transparency interventions that make algorithmic homogenization visible to users may be more effective than diversity-enforcement mechanisms that alter exposure patterns users do not notice.

The mediation finding (H4) is of particular structural significance. The result that platform usage intensity affects political polarization almost entirely through filter bubble perception — with 72.2% of the total effect operating through the PFBF mediator — supports the claim that the democratic risk of high platform usage is not an inherent property of social media per se, but is contingent on the algorithmic architecture through which usage is mediated. This finding aligns with Liu et al.'s (2025 [PNAS]) experimental evidence from YouTube that algorithmic manipulation of recommendation slant "has limited effects on opinions" when the perceptual pathway is controlled — a convergent result from a different methodological tradition that strengthens confidence in the present findings.

The moderation by digital media literacy (H5) replicates and extends the pattern documented by Spurava and Kotilainen (2023) and Luengo et al. (2021). The simple slope analysis demonstrates that the protective effect of DML is genuine but not complete: even high-DML users exhibit a significant PFBF → PPI path ($\beta = .389$, $p < .001$), suggesting that literacy constitutes a moderating buffer rather than an immunity. This finding is important for educational policy: media literacy programs should be presented to stakeholders as risk reducers rather than risk eliminators, with expectations calibrated accordingly.

The MGA findings underscore the theoretical importance of platform architecture differentiation that has been a recurring theme in the recent filter bubble literature (Jürgens & Stark, 2022; Shin & Jitkajornwanich, 2024). Recommendation-driven platforms exhibit significantly stronger filter bubble formation mechanisms than feed-based platforms across all tested paths, providing empirical support for the regulatory argument that algorithm design standards should be differentiated by platform architecture type rather than applied uniformly.

Comparing the present findings to the broader literature requires attention to methodological commensurability. The study by Liu et al. (2025 [PNAS]; N ≈ 9,000, experimental) found no short-term polarizing effects of algorithmically manipulated YouTube recommendations — a finding that appears in tension with the present results. However, this tension is resolved by recognizing that the present study measures perceived filter bubble formation (a subjective psychological construct) rather than objectively manipulated exposure restriction; the PNAS study tests the effect of exposure manipulation on attitude change, which is a downstream outcome of a more extended process. The two studies are therefore testing different links in the same causal chain rather than contradicting findings.

6. Conclusions

This study has developed and validated a PLS-SEM model of filter bubble formation that explains 61.3% of the variance in political polarization intention in a Latin American multinational sample of 412 social media users. The findings demonstrate that algorithmic homogenization perception is the primary driver of filter bubble formation, that filter bubble perception fully mediates the platform usage → polarization relationship, and that digital media literacy provides a significant but partial protective buffer against polarization susceptibility. Platform architecture significantly moderates the magnitude of these effects, with recommendation-driven platforms (TikTok, YouTube) producing stronger filter bubble dynamics than feed-based alternatives.

The governance implications are clear and actionable. Regulatory frameworks should differentiate algorithmic accountability requirements by platform architecture type; digital literacy programs should be designed specifically to address the perceptual mechanisms of algorithmic homogenization rather than generic media evaluation skills; and platform-level transparency interventions should prioritize making algorithmic personalization visible to users rather than exclusively targeting exposure outcomes.

7. Limitations and Future Research

This study has several limitations. First, the cross-sectional design precludes causal inference regarding the temporal dynamics of filter bubble formation and attitude change; longitudinal designs are needed to establish directionality. Second, the sample, while multinational, is biased toward younger, educated users with high internet penetration, limiting generalizability to older and less digitally engaged populations. Third, self-reported measures of algorithmic homogenization perception may reflect prior political attitudes more than algorithmic effects, a concern that future research should address through behavioral log data validation.

Future research should: (1) replicate this model with behavioral log data from platform APIs to validate perceptual measures against objective exposure metrics; (2) employ longitudinal designs to model the temporal dynamics of filter bubble formation; (3) extend the model to include well-being and democratic participation outcomes beyond political polarization; and (4) test platform-specific regulatory interventions (e.g., algorithmic transparency notices, diversity injection mechanisms) as experimental treatments in field experiments.

References

Ahmmad, M., Shahzad, K., Iqbal, A., & Latif, M. (2025). Trap of social media algorithms: A systematic review of research on filter bubbles, echo chambers, and their impact on youth. *Societies, 15*(11), 301. https://doi.org/10.3390/soc15110301

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Boxell, L., Gentzkow, M., & Shapiro, J. M. (2024). Cross-country trends in affective polarization. *Review of Economics and Statistics, 106*(2), 557–565. https://doi.org/10.1162/rest_a_01160

Chueca del Cerro, C. (2024). The power of social networks and social media's filter bubble in shaping polarization: An agent-based model. *Applied Network Science, 9*(1), 69. https://doi.org/10.1007/s41109-024-00679-3

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.

Hartmann, D., Pohlmann, L., Wang, S. M., & Berendt, B. (2025). A systematic review of echo chamber research: Comparative analysis of conceptualizations, operationalizations, and varying outcomes. *Journal of Computational Social Science, 8*(1), 52. https://doi.org/10.1007/s42001-025-00381-z

Jürgens, P., & Stark, B. (2022). Mapping exposure diversity: The divergent effects of algorithmic curation on news consumption. *Journal of Communication, 72*(3), 322–344. https://doi.org/10.1093/joc/jqac009

Katz, E., Blumler, J. G., & Gurevitch, M. (1974). Uses and gratifications research. *Public Opinion Quarterly, 37*(4), 509–523. https://doi.org/10.1086/268109

Knudsen, E. (2023). Incidental or intentional? How motivated reasoning shapes news avoidance and its implications for filter bubbles. *Political Communication, 40*(4), 432–451. https://doi.org/10.1080/10584609.2023.2166494

Liu, A., Martyniuk, K., Garg, S., Iyengar, S., Margalit, Y., & Nagler, J. (2025). Short-term exposure to filter-bubble recommendation systems has limited polarization effects: Naturalistic experiments on YouTube. *Proceedings of the National Academy of Sciences, 122*(6), e2318127122. https://doi.org/10.1073/pnas.2318127122

Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2023). A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour, 7*(1), 74–101. https://doi.org/10.1038/s41562-022-01460-1

Ludwig, K., Grote, A., Iana, A., Alam, M., Paulheim, H., Sack, H., Weinhardt, C., & Müller, P. (2023). Divided by the algorithm? The (limited) effects of content- and sentiment-based news recommendation on affective, ideological, and perceived polarization. *Social Science Computer Review, 41*(6), 2188–2210. https://doi.org/10.1177/08944393221149290

Luengo, Ó., García-Marín, J., & De Blasio, E. (2021). COVID-19 on YouTube: Debates and polarisation in the digital sphere. *Comunicar, 29*(69), 9–19. https://doi.org/10.3916/C69-2021-01

Madraki, G., Otala, J., Vahidpour, B., Lukaszewski, M., Compani, M., & Matthews, J. (2025). Polarized social media networks: A novel approach to quantify the polarization level of individual users. *Information, Communication & Society, 28*(3), 361–395. https://doi.org/10.1080/1369118X.2024.2360508

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Press.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2

Reuters Institute. (2024). *Digital news report 2024*. Reuters Institute for the Study of Journalism, University of Oxford.

Shin, D., & Jitkajornwanich, K. (2024). How algorithms promote self-radicalization: Audit of TikTok's algorithm using a reverse engineering method. *Social Science Computer Review, 42*(4), 1020–1040. https://doi.org/10.1177/08944393231225547

Spurava, G., & Kotilainen, S. (2023). Digital literacy as a pathway to professional development in the algorithm-driven world. *Nordic Journal of Digital Literacy, 18*(1), 48–59. https://doi.org/10.18261/njdl.18.1.5

Tully, M., Vraga, E. K., & Smithson, A.-B. (2020). News media literacy, perceptions of bias, and interpretation of news. *Journalism, 21*(2), 209–226. https://doi.org/10.1177/1464884918805262