



The measurement and influence factors of social consensus in the context of public events

Ruiyang Chen, Lu An, Yajing Zheng, Gang Li, and Chuanming Yu

DOI: <https://doi.org/10.47989/ir31iConf64114>

Abstract

Introduction. This study examines the formation of social consensus during public events by introducing measurable indicators and identifying key influencing factors.

Method. Social consensus was defined and measured in terms of its breadth and strength. Factors influencing the formation of social consensus was identified. The SHAP interpretation method was applied to analyse key drivers of social consensus formation, followed by a configuration analysis of core influencing factors using fsQCA to identify the configuration path to achieve consensus.

Analysis. A dataset of 155,740 posts was collected from Weibo. Four machine learning regression models were trained and evaluated. The best-performing model was further interpreted using SHAP to identify factors' contributions. FsQCA revealed three distinct pathways to high consensus and two pathways to low consensus.

Results. Consensus is negatively associated with the proportion of individuals with higher education, age difference, and the proportion of individuals with firm opinions, while positively associated with social network structure similarity. Groups that achieve high consensus levels have three key characteristics.

Conclusion(s). The study provides a systematic framework for quantifying and analysing social consensus in public events. The findings clarify both consistent predictors and context-specific mechanisms, offering practical insights for guiding public opinion and managing polarization during crises.

Introduction

In recent years, public events have posed significant threats to global welfare and social safety. Social media has become a critical arena where public opinion forms, polarizes, and influences crisis management. While online discussions can support collective coping (Levine & Gazit, 2025), severe polarization may exacerbate hostility and hinder effective responses (Park, 2021). Individuals may hold different opinions and engage in extensive discussions on social media. Thus, understanding how social consensus emerges is a pressing research challenge.

Social consensus refers to the degree of shared understanding that enables coordinated action. The formation of social consensus is dynamic, shaped by the interaction and evolution of individual viewpoints. Previous studies show that individuals tend to avoid conflict by delivering ambiguous message which may facilitate consensus (Roos et al., 2020). But social consensus formation is influenced by multiple factors, such as diverse identities, communication styles, and information environments. To address these gaps, this study aims to (a) introduce the concepts of breadth of consensus and strength of consensus, (b) propose a method to measure the degree of social consensus, and (c) identify factors and mechanisms influencing consensus. The findings aim to precisely assess public opinion risks and to prevent extreme situations.

Grounded in the social identity theory (Tajfel et al., 1971) and the information ecology theory (Nardi & O'Day, 1999), this study addresses the following research questions:

RQ1: How to identify and quantify social consensus in the context of public event?

RQ2: Which factors influence the formation of social consensus?

RQ3: How do the key factors influence the formation of social consensus in the context of public event?

Literature review

Theoretical background

Social identity theory provides an important aspect for understanding consensus formation. Social identity theory suggests that individuals define themselves through social groups based on attributes such as interests, location, values, or relationships, which in turn shape preferences and biases. Huddy and Khatib (2007) found that social identity influenced political decisions, with individuals often voting to align with group expectations. Asch (1951) found that when minority opinions diverge from the majority, individuals often align themselves with the group to reduce conflict, a phenomenon known as the 'conformity effect'. This tendency shows that consensus is not solely an outcome of rational deliberation, but also a product of social pressures and identity-based alignment.

The theory of Information ecology shifts the focus from individuals to the broader systems in which they operate. First proposed by Capurro in 1990, the concept introduces information as embedded in contexts of language, interpretation, and power. Capurro further introduced the notion of information pollution, highlighting relations between information abundance and scarcity (Capurro, 1990). Building on this, Nardi and O'Day (1999) extended the idea of information ecosystem to more practical framework composed of people, practices, technology, and values. In this framework, consensus can be studied through the interplay of information content, information agent, and information environment.

Research on the social consensus

Consensus is the shared understanding or agreement among group members, often measured by the degree of consistency in viewpoints. In the process of reaching consensus, consensus can be interpreted in different ways, from strict consensus to soft consensus (Tran et al., 2024). It can also

be defined in binary terms, from complete agreement (1) to complete disagreement (0) (Naroll, 1983). This type of consensus may be difficult or even impossible to achieve, especially in large and diverse groups. Thus, Elzinga et al. (2011) proposed the concept of soft consensus, allowing the consensus degree to range between 0 and 1, and setting consensus degree thresholds (such as 70%, 80%, etc.) to indicate consensus degree.

Consensus emerges as mutual agreement satisfying all the group members (Liang et al., 2021; Rothstein & Butler, 1987). Users act more readily when content is endorsed by peers (Roth et al., 2021). The theory of communicative action by Habermas and McCarthy (1981) views the public sphere as an arena for dialogue, where rational communication facilitates shared understanding and consensus to some extent. Thus, social consensus can be defined as a shared understanding achieved through social interaction and practice on a common topic. With the prevalence of social media, the scope of consensus has expanded. There is no fixed regulations, necessitating further research into consensus mechanism in social networks (Wu, 2024).

Rawls (1987) proposed the idea of overlapping consensus based on public reason where individuals, setting aside self-interest and social identity, engaged in fair and rational discussions to reach a degree of consensus. The advent of social media has prompted scholars to focus on online consensus, where echo chambers and false consensus effects reinforce existing views and biases, often leading to group polarization (Bakshy et al., 2015; Barberá et al., 2015). According to the group polarization theory, echo chambers can reinforce existing views within a group, pushing the group as a whole toward extreme positions (Sunstein, 2002).

During public events, social media provide abundant information for researchers so that they can conduct interdisciplinary research to enrich studies on consensus. Stevens (1966) measured opinion strength and social consensus using psychophysical methods. Bjornsdottir et al. (2021) found that perceiver consensus reduced biases and improved accuracy. Wang et al. (2021) revealed that social consensus positively affected moral decision-making. Predictive models were also used to study consensus. Babaei et al. (2018) built a classification prediction model based on publisher and tweet characteristics to identify and label high-consensus news tweets. For group decision-making in emergency scenarios, Chen et al. (2022) developed a group consensus model.

In summary, researchers have conducted some research on social consensus, covering different aspects such as theoretical frameworks of social consensus, false consensus, and group decision-making consensus in public events. Few studies systematically analyse the formation mechanism of consensus on social media in the context of public events.

Research on consensus formation mechanism

The evolution of public opinion complicates consensus formation. Since the concept of opinion dynamics was introduced in 1956, researchers have developed various models with distinct evolution rules, such as the DeGroot model, the Friedkin-Johnsen model, and the bounded confidence model. However, few studies have been done to reveal the consensus formation mechanism in the context of public events.

Some scholars have investigated factors influencing consensus formation. For example, Na and Duckitt (2003) found that age and gender significantly affected the level of consensus. Schwartz and Sagie (2000) demonstrated that regional economic development and democratization had notable impact on social consensus formation. Dong et al. (2017) examined the influence of the number of opinion leaders in social networks on consensus formation using the DeGroot model. Zhou et al. (2020) introduced self-persistence to measure individuals' adherence to initial opinions and analysed its influence on consensus. Tran et al. (2024) used the F-J model, highlighted differences between consensus formation on a sequence of issues and that on individual issues. Zhang et al. (2022) employed a weighted opinion dynamics model and showed that trust

relationships and opinion similarity between individuals positively influenced consensus formation. Dong et al. (2022) demonstrated that similarity in opinions and social network structures enhanced the efficiency of reaching consensus. Additionally, Liu and Tse (2014) observed that in highly connected communities within social networks, consensus formed quickly within communities but slowly between them.

In conclusion, existing research has explored factors of social consensus formation, such as gender, age, and regional economic development. Many studies rely on simulations with opinion dynamics models. More research needs to be conducted using real social media data for opinion mining and consensus identification. Most studies focus on individual factors like personal traits, social network structures, trust, or opinion similarity, with few studies systematically examining the roles of information agents, information content, and the information environment. Additionally, previous research often examines single factors, neglecting the combined effects of multiple factors and overlooking causal asymmetry.

Methods

This study explores the mechanism of social consensus formation in the context of public events. The overall research framework is shown in Figure 1.

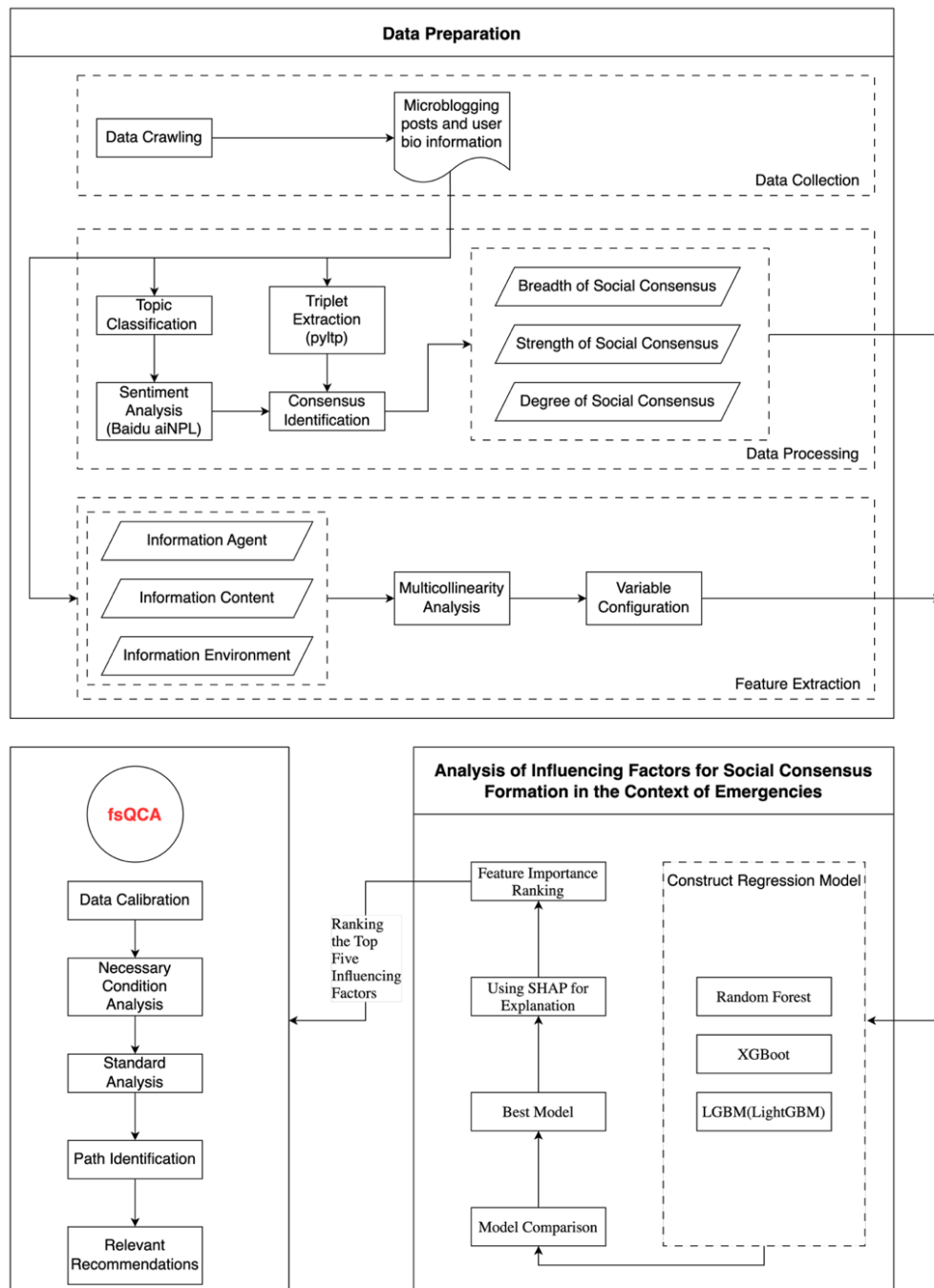


Figure 1. Research framework of the mechanism of social consensus.

Identification of social consensus

Sentiment analysis is a critical component of opinion identification in the context of public events. The Baidu aiNLP sentiment analysis system integrates lexicon-based methods with deep learning models and has demonstrated strong generalization ability and accuracy in Chinese-language sentiment classification. Thus, the Baidu aiNLP tool was employed to analyse the sentiment polarity of microblogging posts in this study. Pytlp is an open-source Chinese NLP toolkit that provides various functionalities to support fine-grained viewpoint extraction.

A three-step method was employed to identify social consensus. First, topics of posts were recognized by the Word2Vec model and K-means clustering. Sentiment analysis was then conducted using the Baidu aiNLP tool, which classifies sentiment polarity as negative (0), neutral (1), or positive (2) and provides confidence scores for reliability of sentiment classification results. Finally, fine-grained viewpoint recognition was achieved by extracting Subject-Predicate-Object (SPO) triplets using the PyItp toolkit. Semantic role annotation and dependency syntax methods are applied to extract subjects, predicates, objects, and their relationships, forming triplet representations for posts.

Measurement of social consensus

In this study, social consensus is understood as a group-level phenomenon in which individuals converge toward a shared viewpoint within a discussion context. Rather than treating it as a purely binary outcome, we conceptualize social consensus as a continuous phenomenon that varies in degree across groups.

Consensus degree is measured along two dimensions, breadth, and strength. Achieving consensus relies on viewpoint consistency and emotional resonance. Building on these two dimensions, this study proposes a method to measure social consensus by integrating opinion and emotional consistency.

Breadth of social consensus is defined as the proportion of individuals who share the mainstream viewpoint on a given topic at a given time. The attitudes of individuals can be simplified as positive, negative, and neutral.

Strength of social consensus reflects to the extent of opinion alignment among users who share the same viewpoint on a given topic at a given time. It is measured by combining opinion consistency and sentiment consistency. Opinion consistency is represented by the similarity of SPO triplets extracted from the group, while sentiment consistency is represented by the similarity of emotional expression within the group.

Following Boltzmann's classical formulation in physics, where extensive magnitudes are expressed as the product of breadth and strength, we analogously define social consensus as the product of its breadth and strength, normalized through a max-min procedure.

Factors influencing the formation of social consensus in the context of public events

Drawing on the social identity theory and information ecology theory, fifteen factors were identified across three dimensions: information agent, information content, and information environment.

Information agent

Social identity theory suggests that individuals form social groups based on attributes like interests, location, and relationships. Shared group identity increases value alignment and facilitates consensus on public-event topics (Jian and Zhang, 2025). Beyond basic demographics such as gender, age, industry, education, and location, social media users exhibit characteristics like popularity, stakeholder roles, self-persistence, and opinion leadership. These attributes significantly influence the formation of social consensus. Among these factors, we measured popularity via the follower to following ratio (FF ratio) (Ji et al., 2023). Referring to the classification method of stakeholders (An et al., 2021), stakeholders types are inferred from authentication information, profile, and industry tags of microblogging users. Self-persistence was measured as the degree of stability between pre- and post-event viewpoints, while opinion leadership was derived from repost, comment, like, and follower counts using an entropy-weighted composite

score. These indicators represent how users' social identity and influence affect consensus formation (See Table 1).

Influencing factor	Description	Value
Gender	Male/female	0,1
Age	Under 18 / 19-35 / 36-60 / Above 60	1,2,3,4
Location	Economic development level of the region	1,2,3,4,5,6
	Geographical position of the region	1,2,3,4,5
Education level	Primary school / Junior high school / Senior high school (including vocational) / College and above	1,2,3,4
Industry	Culture & Arts / Film & Entertainment / Finance & Insurance / IT / Manufacturing / Media & PR / Healthcare / Education & Research / Business Services / Public Affairs / Legal / Others	1,2,3,4,5,6,7,8,9,10,11,12
Popularity	Ration of the number of followers to the number of following	[0, ∞]
Stakeholders	Government departments / Legal institutions / Legal professionals / Mainstream media	1,2,3,4,5,6,7
Self-persistence	Opinion holder / Opinion neutral / Opinion tolerant	1,2,3
Opinion leader	Whether the user is an opinion leader	0,1

Table 1. Influencing factors of information agent

Information content

Grounded in persuasion and emotional contagion theories, user discussions on specific topics influence how others perceive and adjust their own views, thereby accelerating or delaying the process of consensus formation.

This study categorizes information features into four dimensions. The theme of a post represents the core topical focus. Emotion refers to the affective tone conveyed in the posts. Form refers to how information is presented. Time reflects when the information is released. Collectively, these features illustrate how the attributes of information content contribute to the development of social consensus (See Table 2).

Influencing factor	Description	Value
Topic	The type of topic included in the microblog post	[1, K], where K is the number of topics
Sentiment	The sentiment tendency of the microblog post	0,1,2
Length	The number of characters in the microblog post	1,2,3
Content format	Whether the microblog post includes images, videos, or links	0,1
Time of posting	The time at which the microblog post was published	1,2,3

Table 2. Influencing factors of information content

Information environment

Individuals engage in information environments defined by trust and network structures. Interpersonal trust emerges from observable interactions like reposting, liking, and commenting. The trust levels can be quantified by the strength of interaction between users, which plays a crucial role in promoting consensus formation (Sagi & Rokach, 2021). Users receive information from neighbour nodes and adjust their opinions based on their current views and the information

they encounter. Higher structural similarity increases the likelihood of connections, exchange of views, and eventual consensus (Dong et al., 2022) (See Table 3).

Influencing factor	Description	Value
Trust between Individuals	The trust level of individuals participating in the discussion of topic t at time d	0,1
Similarity of network Structure	The similarity of network structure among individuals discussing topic t at time d	[0,1]

Table 3. Influencing factors of information environment

Influence mechanism of the social consensus formation in the context of public events

To identify the key factors influencing social consensus, four machine learning regression models were constructed: XGBoostRegressor, LGBMRegressor (Prokhorenkova et al., 2018), CatBoostRegressor (Haneczok & Piskorski, 2020), and RFRegressor (Lundberg & Lee, 2017). Model performance was assessed using three indicators: MAE, MSE, and R^2 .

Analysis of influencing factors based on SHAP interpretation

To understand the contribution of each feature to consensus prediction, we employ the Shapley Additive explanations (SHAP) model to interpret features' contributions.

To evaluate the combined effects of multiple factors, the fsQCA technique was employed to explore the configurational mechanism influencing social consensus formation, using the most important five factors identified by the SHAP model.

The analysis process includes four steps: data calibration, necessity analysis of individual conditions, sufficiency analysis of condition combinations, and robustness testing. During the data calibration, antecedent variables were standardized to ensure comparability across dimensions, and a fuzzy set membership score matrix was constructed. The direct calibration method, which assigns values of 1 (full membership), 0.5 (crossover point), and 0 (full non-membership) (Ragin, 2001). Anchor points were determined by the distribution of the raw data, typically using percentiles such as 25%, 50%, and 75%.

The necessity of single conditions was then evaluated by testing their consistency exceeded the conventional threshold of 0.9 while maintaining sufficiently high coverage. Consistency and coverage were calculated from fuzzy membership scores. In the sufficiency analysis phase, a truth table was generated to examine combinations of antecedent conditions leading to the outcome. A frequency threshold was set to exclude configurations supported by only a small number of cases. Thresholds for raw consistency and Proportional Reduction in Inconsistency (PRI) were set and the values in the truth table were manually altered.

Finally, robustness of the results was assessed by varying calibration anchors, adjusting case frequency thresholds, and modifying PRI consistency thresholds.

Analysis and Results

Data collection and pre-processing

This study investigated the event of restaurant attack in Tangshan in 2022. The incident sparked widespread debates on public safety, gender violence, and social justice. We collected 155,740 posts related to the incident on Weibo from June 10 to September 30, 2022. After deduplication and cleaning, 129,553 unique posts and 91,960 users' profiles remained, including attributes such as

gender, age, location, and social network information. Texts were pre-processed through removal of irrelevant symbols (e.g., @, #, URLs), word segmented using Jieba, and stop-word filtering with a refined dictionary of 3,883 terms.

Topical analysis

The pre-processed posts were trained with a 25-layer Word2Vec model. The K-means clustering was then applied to group the posts into 11 topics. Sentiment analysis with Baidu AI indicated that 90.83% of posts expressed negative sentiment. To enable structural analysis, semantic role labelling and dependency parsing extracted 507,934 SPO triples, with 298,808 retained after cleaning.

User information analysis

A total of 91,960 users were analysed. Based on the opinion evolution model, users were further classified into firm, neutral, and tolerant groups, with firm users being the most common (5,738 cases). Opinion leaders were identified using the entropy-weighted indicators of reposts, comments, likes, and followers, and the top 1% of high-influence users (7,130) were identified as opinion leaders.

Experimental data preparation

Consensus subjects, defined as groups discussing the same topic, were segmented by topic consistency and information release time to account for differences in information environments. Group characteristics, i.e., information agent, information content, and environment were analysed.

The consensus degree was calculated based on emotional and viewpoint consistency. The experimental dataset was grouped by topic and posting time. A total of 2,600 groups with at least two posts were selected as experimental samples for analysis. Control variables included theme categories (t) and date and time of the post (d).

Independent variables were derived from three dimensions: agent (i.e., gender, age, education, popularity, firm views, opinion leaders), content (i.e., sentiment ratio, short/pure text ratio), and environment (i.e., trust, network similarity). Independent variables were listed in Table 4.

Variables		Description
A1	female_ratio	Proportion of females in the group
A2	age_gap	Age standard deviation in the group
A3	university_ratio	Proportion of individuals with higher education in the group
A4	income_gini	Income gap of individuals within a group
A5	location_gini	The degree of dispersion in the spatial distribution of individuals within a group
A6	industry_gini	The dispersion degree of individual industry distribution in the group
A7	FF_avg	Average popularity of individuals within a group
A8	p_ratio	Proportion of ordinary people in the group
A9	spd_ratio	Proportion of individuals with firm opinions in the group
A10	opl_ratio	Proportion of opinion leaders in the group
B1	positive_ratio	The proportion of individuals in the group whose posts have positive emotions
B2	st_ratio	The proportion of individuals whose posts are short text in the group
B3	pt_ratio	The proportion of individuals who post pure text in the group
C1	trustdegree_avg	The average level of trust among individuals within a group
C2	networksim_avg	The average network structure similarity among individuals in a group

Table 4. Independent variables of social consensus prediction model for public events.

The dependent variable was the degree of social consensus. We conducted sentiment analysis on the posts by individuals in each group, and divided them into three opinion camps, including supportive, neutral, and opposing, and the majority camp was identified as the mainstream opinion. Consensus breadth was calculated by the as the proportion of individuals in the mainstream camp. Consensus strength was calculated by emotional and viewpoint similarity. The overall consensus degree of each group was thus obtained by combining the breadth and strength of consensus.

Analysis of Factors Influencing the Formation of Social Consensus in the context of public events

Machine learning regression model

The data was first pre-processed by filling missing values with the median or mean by winsorizing extreme values. To address multicollinearity, variance inflation factors (VIFs) were calculated. Following (O'Brien, 2007), the variable with VIFs above 10 (trustdegree_avg) was removed.

The processed dataset was then used to train four models, and their performance was evaluated using MAE, MSE, and R^2 . As shown in Table 5, the CATBoost model demonstrated the best performance in terms of MSE and R^2 and was therefore selected as the optimal model for predicting consensus degree.

Model	Performance on the training set			Performance on the training set		
	MAE	MSE	R^2	MAE	MSE	R^2
XGBoostRegressor	17.7518	1468.8421	0.5979	17.9859	1502.6846	0.5855
CatBoostRegressor	17.9151	1098.3819	0.7037	18.5254	1176.9550	0.6753
LGBMRegressor	19.0151	1234.1868	0.5547	19.5629	1305.4494	0.5289
RFRegressor	19.4738	1364.8115	0.5076	19.9061	1417.2267	0.4885

Table 5. Performance of different models.

SHAP-based interpretation

The experimental data was trained using the CATBoost model. SHAP values were employed to interpret the contribution of each feature to consensus predictions. Global feature summary (See Figure 2) and feature importance plots (See Figure 3) were generated.

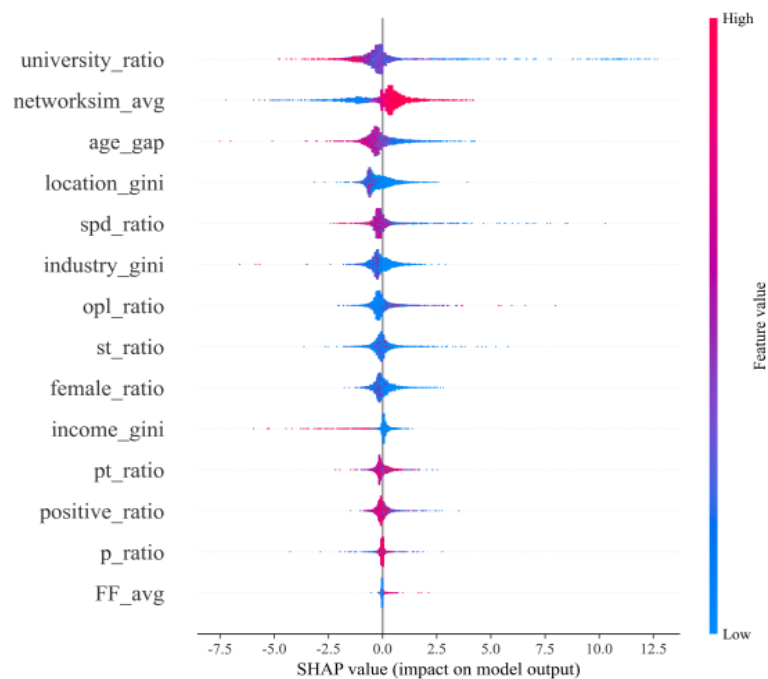


Figure 2. Summary plot of the influence of features on consensus.

Results showed that consensus was negatively associated with the proportion of highly educated individuals, age heterogeneity, and firm opinion holders, but positively associated with network structure similarity. Spatial dispersion showed no significant effect. Higher education introduced diversity of perspectives, age heterogeneity reflected generational differences, and firm views reduced flexibility, all of which impeded consensus. Conversely, structural similarity in social networks facilitated connections and opinion exchange, promoting consensus.

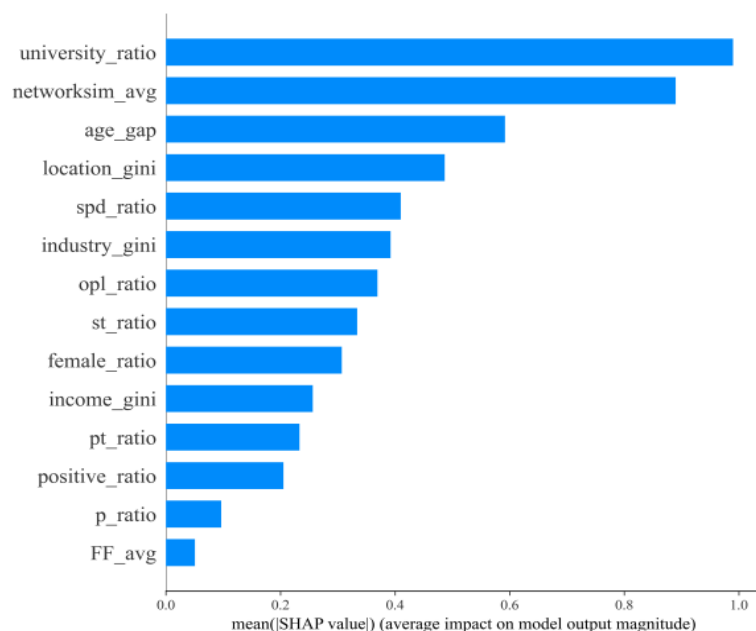


Figure 3. Feature importance of factors influencing social consensus in the context of public events.

Topic-specific analysis

The SHAP global feature analysis for the 11 topics shows the magnitude and direction of each variable's influence on consensus. Higher education and network similarity remained the top factors, while the proportion of firm opinion holders and income inequality typically appeared in the mid-level ranks.

Specific topics revealed stronger effects of other variables. In the topic of 'official information release' (Figure 4), firm opinion holders became the most influential, highlighting how untimely or inaccurate communication fosters distrust and amplifies scepticism, a dynamic often described as the 'Tacitus trap'. In the topic of 'women's protection' (Figure 5), income inequality ranked the highest, as socioeconomic differences led to divergent priorities. High-income groups stressed gender equality, while lower-income groups emphasised survival needs.

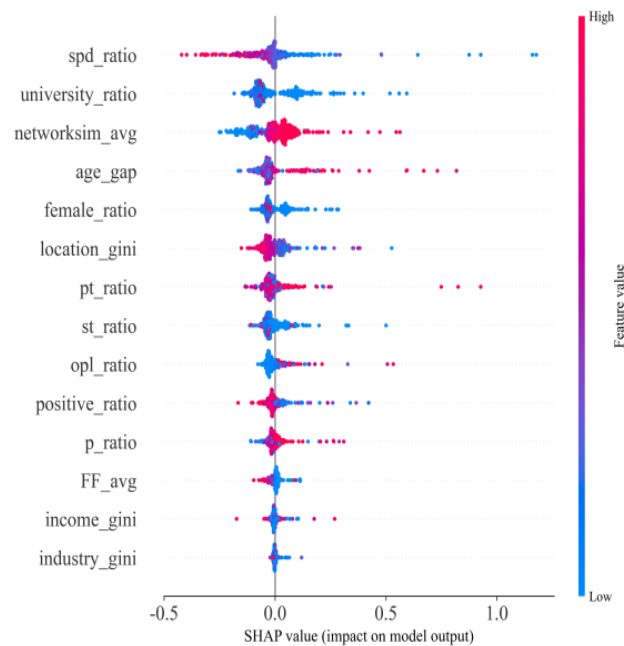


Figure 4. SHAP global feature summary (official information release).

These findings underscore that consensus mechanisms are not uniform but shaped by the interplay between structural attributes and thematic content.

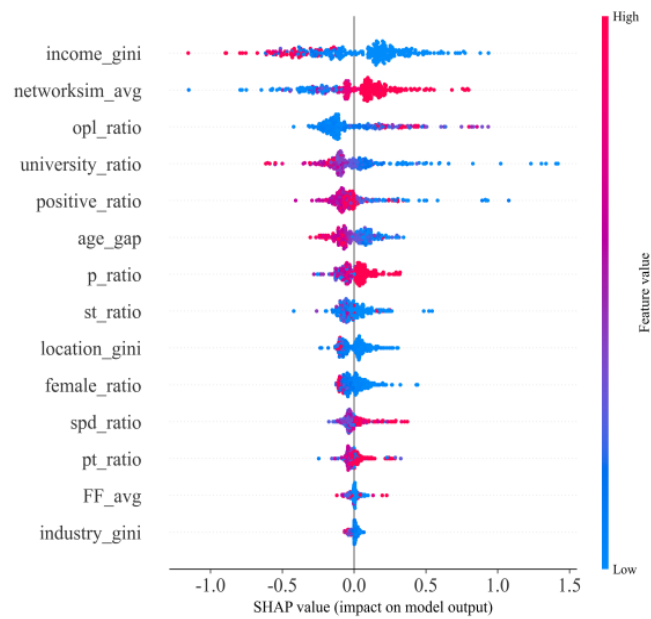


Figure 5. SHAP global feature summary (women's protection).

Breadth and strength of consensus

This study analyses the key factors influencing breadth and strength of consensus to explore the formation mechanism of social consensus in the context of public events. For the breadth of consensus, the XGBoost model achieved the best performance ($R^2 = 0.6277$). SHAP values were then applied to interpret the results (Figure 6). For the strength of consensus, the CATBoost model performed the best ($R^2 = 0.6567$), and SHAP interpretation was again employed (Figure 7).

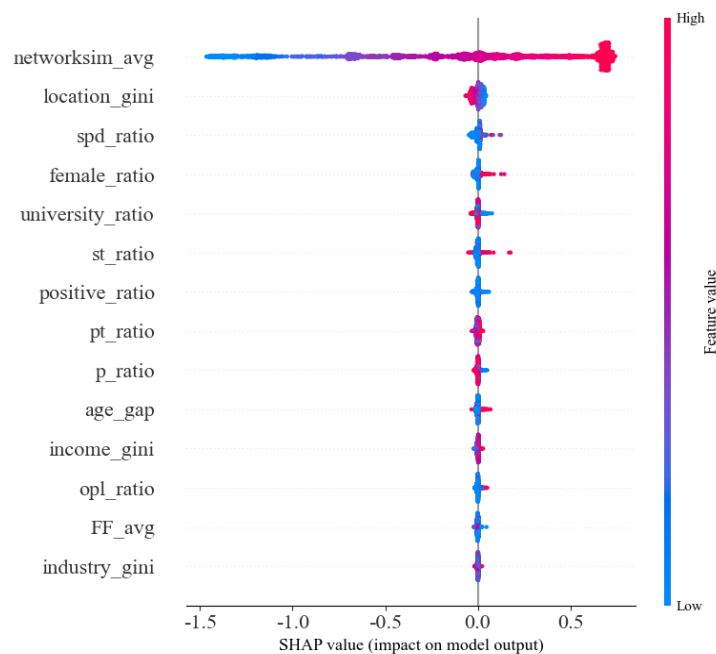


Figure 6. Global feature summary of consensus breadth based on the SHAP values.

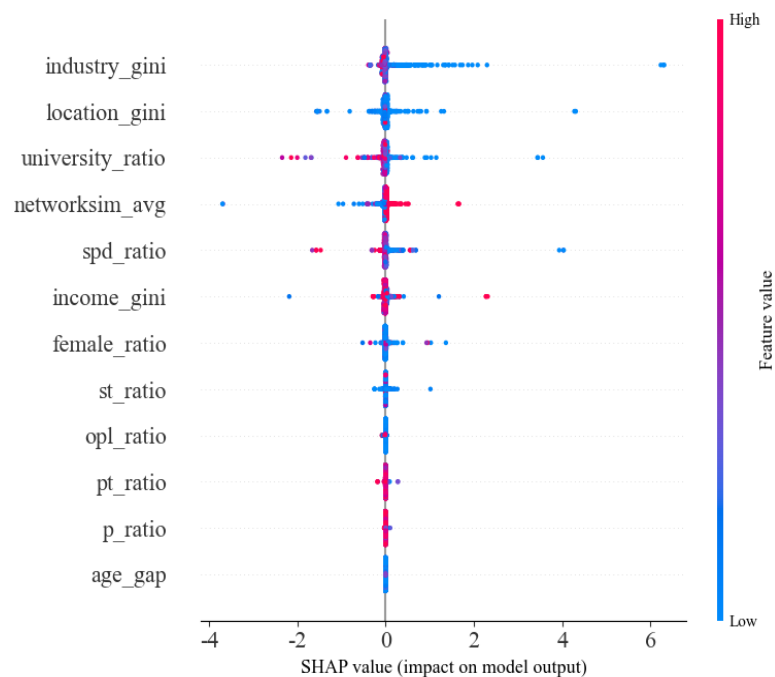


Figure 7. Global feature summary of consensus strength based on the SHAP values.

A comparing analysis of the top five factors influencing consensus degree, breadth, and strength is shown in Table 6. Two results are consistent across degree, breadth, and strength: higher-education ratio is significantly negative and network similarity is significantly positive. Several differences clarify mechanism. The share of firm-view users is significantly negative for consensus degree and strength but not for breadth, indicating that firm stances do not prevent initial alignment around a mainstream stance but impede deeper convergence within that camp. Spatial dispersion lowers consensus breadth but not strength, suggesting geographic proximity extends the reach of alignment without necessarily intensifying within-camp cohesion. The female ratio is positively associated with consensus breadth but unrelated to strength or overall degree, consistent with greater openness to external influence extending the aligned camp without deepening within-camp similarity. Industry dispersion reduces consensus strength but not breadth or overall degree, implying cross-professional barriers hinder fine-grained convergence. Age heterogeneity does not directly affect consensus breadth or strength but lowers overall degree.

Independent variables		Dependent variables		
		Consensus degree	Breadth of consensus	Strength of consensus
university_ratio	Proportion of individuals with higher education in the group	▲	▲	▲
networksim_avg	The average social network structure similarity	■	■	■
spd_ratio	Proportion of individuals with firm opinions	▲	●	▲
location_gini	Individual spatial dispersion	●	▲	●
female_ratio	Proportion of females	●	■	●
industry_gini	Individual industry gap	●	●	▲
age_gap	Age standard deviation	▲	●	●

Table 6. Comparative analysis of factors influencing the breadth and strength of consensus

Note: ▲ indicates a significant negative correlation between the independent variable and the dependent variable. ■ indicates a significant positive correlation between the independent variable and the dependent variable. ● indicates no significant correlation between the independent variable and the dependent variable.

Configuration mechanism of core factors influencing the formation of social consensus

Data calibration

The fsQCA method, grounded in Boolean algebra, analyses membership scores on a continuous scale from 0 to 1. Following Lewellyn and Muller-Kahle (2022), the 10th, 50th, and 90th percentiles of causal and outcome variables were set as anchors representing non-membership, the crossover point, and full membership, respectively. To enhance precision, membership scores were reduced by 0.01 to ensure cases with a score of 0.5 are included in the fuzzy set analysis. See Table 7.

Variable type	Variable name	Variable code	Description
Independent variable	age_gap	A2	Age standard deviation in the group
	university_ratio	A3	Proportion of users with higher education in the group
	location_gini	A5	The degree of dispersion in the spatial distribution of individuals within a group
	spd_ratio	A9	Proportion of individuals with firm opinions
	networksim_avg	C2	The average network structure similarity among individuals in a group
Dependent variable	Consensus degree	SCL	The consensus reached by the group on topic <i>t</i>

Table 7. Basic information of the variables in the FSQCA model.

Necessity analysis of individual conditions

In fsQCA, a condition is considered necessary when its consistency exceeds 0.9 (Rihoux & Ragin, 2009). Using the necessity analysis function of the fsQCA tool, both the presence and absence of individual conditions were analysed. The results indicate that all the conditions have consistency below 0.9, suggesting that no single condition is necessary to enhance consensus.

Sufficiency analysis of condition configurations

The QCA technique examines the combined effects of multiple variables, under the assumptions of equifinality, conjunctural causation, and asymmetry. In this study, five core conditions were selected for configurational analysis. Using the fsQCA tool, a truth table was generated based on the fuzzy set membership matrix. The truth table included all logically possible configurations, along with the number of cases, raw consistency scores, and PRI consistency scores. Configurations with fewer than 60 cases were excluded to retain 75% of the sample (Zhang & Du, 2019). Consistency thresholds were set at 0.8 for raw consistency and 0.5 for PRI (Greckhamer et al., 2018). The refined truth table produced complex, intermediate, and parsimonious solutions using the fsQCA 4.0 tool.

Intermediate solutions for high and low consensus (See Table 8 and Table 9 respectively) showed overall consistency above 0.8 and coverage above 0.5, demonstrating strong explanatory power. Robustness tests further confirm the stability of the findings when the frequency threshold was set to 58 (Schneider & Wagemann, 2012) and calibration anchors to 85%, 50%, and 15%. These results suggest that the configurational findings are robust.

Path	Condition	Raw coverage	Unique coverage	Consistency
1	$\sim A2^* \sim A5^* C2^* \sim A9$	0.317421	0.0481102	0.834982
2	$\sim A2^* \sim A3^* C2^* \sim A9$	0.348187	0.0788763	0.874104
3	$\sim A3^* \sim A5^* C2^* \sim A9$	0.317079	0.119064	0.846401
Solution coverage				0.515362
Solution consistency				0.821598

Table 8. Intermediate solution for achieving high consensus.

Path	Condition	Raw coverage	Unique coverage	Consistency
1	$A2^* A3^* \sim C2$	0.33152	0.110879	0.876987
2	$A3^* A5^* A9$	0.411515	0.190874	0.877442
Solution coverage				0.522394
Solution consistency				0.852593

Table 9. Intermediate solution for achieving low consensus.

Identifying the configurational path to achieve high consensus

Table 10 and Table 11 summarize the core and peripheral conditions in the three high-consensus paths and the two low-consensus paths, defining a socially driven consensus mode.

Path S1, a strong-tie socially driven mode, is characterized by significant age difference, high social network similarity, small income gap, and low self-persistence. High trust, frequent interaction, and social pressures reinforce group values and promote consensus.

Dependent variables		Configuration paths		
		S1	S2	S3
A2	age_gap	●	⊗	
A3	university_ratio		⊗	⊗
A5	location_gini	⊗		⊗
A9	spd_ratio	⊗	⊗	●
C2	networksim_avg	●	●	●

Table 10. Configurational paths of high consensus.

Note: ● indicates that the core condition exists. ⊗ indicates that the core condition is missing.

Dependent variables		Configuration paths	
		NS1	NS2
A2	age_gap	●	
A3	university_ratio	●	●
A5	location_gini		●
A9	spd_ratio		●
C2	networksim_avg	⊗	

Table 11. Configurational paths of low consensus.

Note: ● indicates that the core condition exists. ⊗ indicates that the core condition is missing. ● indicates that the edge condition exists.

Path S2, the weak-tie socially driven mode, features small age gap, high social network similarity, low education level, and low self-persistence. Consistent with the weak tie hypothesis and social resource theory (Granovetter, 1977), weak ties provide access to diverse information and broader social support, which significantly contribute to the expansion of social resources. In lower-educated groups, simplified information processing and reliance on trusted sources further contribute to agreement.

Path S3, the firm-opinion socially driven mode, is characterized by high self-persistence, high social network similarity, small income gap, and low education level. While firm opinion holders usually hinder consensus, individuals with similar economic and educational backgrounds resonate with each other, reinforcing shared perspectives and producing collective recognition. The mutual understanding and collective recognition within such groups drive the formation consensus.

In contrast, the low-consensus paths demonstrate divergent dynamics. Path NS1, characterized by large age gap, high education level, and low social network similarity, highlights the difficulties in forming shared perspectives across generational and educational divides. Path NS2, defined by large income gap and high education level, reveals that significant socioeconomic differences and the independent thinking fostered by higher education hinder consensus, leading to persistently low level of consensus.

Discussions

Theoretical Significance

The theoretical significance lies in two aspects. First, it proposes a novel method for identifying and measuring social consensus. Previous research on social consensus mostly remains at the theoretical or speculative level, with limited focus on online public opinion during public events. Building on viewpoint recognition techniques, this study identifies user perspectives on Weibo at both thematic and feature levels, using viewpoint and emotional consistency to measure consensus. By defining consensus degree as the product of breadth and strength, we refine the conceptualization of social consensus and provide measurable indicators, bridging the gap between the abstract theory and concrete metrics while offering a robust tool for empirical analysis in dynamic scenarios like public events.

Second, this study reveals the influencing factors and complex mechanisms underpinning consensus formation. This study fills a theoretical gap by constructing a comprehensive model of the factors influencing social consensus formation. Based on social identity theory and information ecology theory, we construct a comprehensive model encompassing 15 factors across the dimensions of information agents, content, and environment. These variables include gender, age, education level, regional economic development and so forth. We systematically analyse the influence mechanism of these factors, empirically validating their roles and uncovering how specific combinations of factors influence consensus formation. Key findings include the negative impact of age gaps on consensus degree, the delaying role of firm opinions on consensus strength, and the positive effect of network similarity on both breadth and strength. Additionally, this study reveals nuanced relationships, such as how the income gap intensifies negative effects on specific topics like 'women's protection' while a higher proportion of women in a group positively correlates with consensus breadth. While this context provides a clear and analytically tractable setting for examining consensus formation under crisis conditions, it may also shape the relative importance of certain factors. In events with more positive or emotionally balanced sentiment structures, the mechanisms identified here, such as the roles of education heterogeneity, firm opinions, or network similarity, may operate differently or with varying intensity.

Practical Implication

Public events, characterized by suddenness and unpredictability, often provoke widespread public discussion and impact social stability. This study provides a vital theoretical framework and practical guidance for understanding and guiding public opinion during public events. By dynamically identifying consensus and divergent opinions, we could enhance the ability to monitor public opinion and mitigate risks of opinion polarization. Insights on key influencing factors, such as education level, network similarity, and income gap, offer actionable recommendations for fostering broad and deep social consensus. In practice, these findings can help refine the design of public communication strategies and strengthen the responsiveness of governance systems in addressing social concerns.

Conclusions

This study develops methods to measure and explain social consensus in the context of public events. Drawing from social consensus and public opinion studies, social consensus was quantified through strength and breadth of consensus. Based on the social identity theory and information ecology theory, 15 key influencing factors were analysed across information agents, information content, and information environment.

Future research could extend this study by incorporating more research methods and examining multiple types of public events to compare consensus formation across events with diverse emotional valence and to test the generalizability and contextual sensitivity of the revealed

mechanism. For instance, user experiments and simulations could complement current methods. Methodologically, integrating qualitative and quantitative approaches may also help identify the key factors affecting consensus formation. These efforts can help shape insights into the group-level consensus, thus building a more comprehensive theoretical and practical understanding in the context of public events.

Acknowledgements

This study was supported by the National Social Science Foundation of China (Grant No. 23&ZD230).

About the authors

Ruiyang Chen is a PhD student of Information science at the Center for Studies of Information Resources and the School of Information Management, Wuhan University, China. Her research interests include social media analysis and digital trust. She can be reached at rychen22@whu.edu.cn.

Lu An is a professor at School of Information Management and deputy director of the Center for Studies of Information Resources, Wuhan University, China. She is the corresponding author of this paper. Her research focuses on crisis informatics and social media analysis. She can be contacted at anlu97@163.com.

Yajing Zheng is a master at the School of Information Management, Wuhan University, China. She can be contacted at 17362990004@163.com.

Gang Li is a professor and deputy director of the Center for Studies of Information Resources, Wuhan University, China. He can be contacted at ligang@whu.edu.cn.

Chuanming Yu is a professor at the School of Information Engineering, Zhongnan University of Economics and Law, China. He can be contacted at yuchuanming2003@126.com.

References

- An, L., Zhou, W., Ou, M., Li, G., & Wang, X. (2021). Measuring and profiling the topical influence and sentiment contagion of public event stakeholders. *International Journal of Information Management*, 58(7), 102327.
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In *Groups, leadership and men; research in human relations*. (pp. 177-190). Carnegie Press.
- Babaei, M., Kulshrestha, J., Chakraborty, A., Benevenuto, F., Gummadi, K. P., & Weller, A. (2018). *Purple Feed: Identifying High Consensus News Posts on Social Media* Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, New Orleans, LA, USA. <https://doi.org/10.1145/3278721.3278761>
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132. <https://doi.org/10.1126/science.aaa1160>
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychol Sci*, 26(10), 1531-1542. <https://doi.org/10.1177/0956797615594620>

- Bjornsdottir, R. T., Hehman, E., & Human, L. J. (2021). Consensus Enables Accurate Social Judgments. *Social Psychological and Personality Science*, 13(6), 1010-1021. <https://doi.org/10.1177/19485506211047095>
- Capurro, R. (1990). Towards an information ecology. In I. Wormell (Ed.), *Information Quality: Definitions and Dimensions* (pp. 122-139). Taylor Graham.
- Chen, X., Zhang, W., Xu, X., & Cao, W. (2022). Managing Group Confidence and Consensus in Intuitionistic Fuzzy Large Group Decision-Making Based on Social Media Data Mining. *Group Decision and Negotiation*, 31(5), 995-1023. <https://doi.org/10.1007/s10726-022-09787-w>
- Dong, Q., Sheng, Q., Martínez, L., & Zhang, Z. (2022). An adaptive group decision making framework: Individual and local world opinion based opinion dynamics. *INFORMATION FUSION*, 78, 218-231. <https://doi.org/10.1016/j.inffus.2021.09.013>
- Dong, Y., Ding, Z., Martínez, L., & Herrera, F. (2017). Managing consensus based on leadership in opinion dynamics. *INFORMATION SCIENCES*, 397-398, 187-205. <https://doi.org/10.1016/j.ins.2017.02.052>
- Elzinga, C., Wang, H., Lin, Z., & Kumar, Y. (2011). Concordance and consensus. *INFORMATION SCIENCES*, 181(12), 2529-2549. <https://doi.org/10.1016/j.ins.2011.02.001>
- Granovetter, M. S. (1977). The Strength of Weak Ties¹¹This paper originated in discussions with Harrison White, to whom I am indebted for many suggestions and ideas. Earlier drafts were read by Ivan Chase, James Davis, William Michelson, Nancy Lee, Peter Rossi, Charles Tilly, and an anonymous referee; their criticisms resulted in significant improvements. In S. Leinhardt (Ed.), *Social Networks* (pp. 347-367). Academic Press. <https://doi.org/10.1016/B978-0-12-442450-0.50025-0>
- Greckhamer, T., Furnari, S., Fiss, P. C., & Aguilera, R. V. (2018). Studying configurations with qualitative comparative analysis: Best practices in strategy and organization research. *Strategic Organization*, 16(4), 482-495. <https://doi.org/10.1177/1476127018786487>
- Habermas, J., & McCarthy, T. (1981). *The Theory of Communicative Action*.
- Haneczok, J., & Piskorski, J. (2020). Shallow and deep learning for event relatedness classification. *Information Processing & Management*, 57(6), Article 102371. <https://doi.org/10.1016/j.ipm.2020.102371>
- Huddy, L., & Khatib, N. (2007). American Patriotism, National Identity, and Political Involvement. *American Journal of Political Science*, 51(1), 63-77. <http://www.jstor.org/stable/4122906>
- Ji, J., Zhu, Y., & Chao, N. (2023). A comparison of misinformation feature effectiveness across issues and time on Chinese social media. *Information Processing & Management*, 60(2), 103210. <https://doi.org/10.1016/j.ipm.2022.103210>
- Jian, F., & Zhang, P. (2025). A comparative analysis of China-themed books in three ASEAN countries: Implications for resource development and intercultural communication. *Data and Information Management*, 9(2), 100081. <https://doi.org/10.1016/j.dim.2024.100081>
- Levine, D., & Gazit, T. (2025). Unorthodox Information Sources of Coping With the COVID-19 Crisis in the Ultra-Orthodox Society. *Social Science Computer Review*, 43(1). <https://doi.org/10.1177/08944393241246282>

- Lewellyn, K. B., & Muller-Kahle, M. I. (2022). A Configurational Exploration of How Female and Male CEOs Influence Their Compensation. *Journal of Management*, 48(7), 2031-2074, Article 01492063211027225. <https://doi.org/10.1177/01492063211027225>
- Liang, H., Yuan, F., Zhou, Z., & Su, H. (2021). Opinion separation in leader-follower cooperative social networks. *Neurocomputing*, 434, 90-97. <https://doi.org/10.1016/j.neucom.2020.12.079>
- Liu, X. F., & Tse, C. K. (2014). Impact of degree mixing pattern on consensus formation in social networks. *Physica A: Statistical Mechanics and its Applications*, 407, 1-6. <https://doi.org/10.1016/j.physa.2014.03.086>
- Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 30 (NIPS 2017)*,
- Na, E.-Y., & Duckitt, J. (2003). Value Consensus and Diversity between Generations and Genders. *Social Indicators Research*, 62(1), 411-435. <https://doi.org/10.1023/A:1022665705561>
- Nardi, B. A., & O'Day, V. (1999). *Information Ecologies: Using Technology with Heart*. The MIT Press. <https://doi.org/10.7551/mitpress/3767.001.0001>
- Naroll, R. (1983). The moral order : an introduction to the human situation / Raoul Naroll.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673-690. <https://doi.org/10.1007/s11135-006-9018-6>
- Park, Y.-E. (2021). Developing a COVID-19 Crisis Management Strategy Using News Media and Social Media in Big Data Analytics. *Social Science Computer Review*, 40, 1358 - 1375. <https://doi.org/10.1177/08944393211007314>
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 31 (NIPS 2018)*,
- Ragin, C. C. (2001). Fuzzy-Set Social Science.
- Rawls, J. (1987). The Idea of an Overlapping Consensus. *Oxford Journal of Legal Studies*, 7(1), 1-25. <http://www.jstor.org/stable/764257>
- Rihoux, B., & Ragin, C. (2009). *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. SAGE Publications, Inc. <https://doi.org/10.4135/9781452226569>
- Roos, C. A., Koudenburg, N., & Postmes, T. (2020). Online Social Regulation: When Everyday Diplomatic Skills for Harmonious Disagreement Break Down. *Journal of Computer-Mediated Communication*, 25(6), 382-401. <https://doi.org/10.1093/jcmc/zmaa011>
- Roth, R., Ajithkumar, P., Natarajan, G., Achuthan, K., Moon, P., Zinzow, H., & Madathil, K. C. (2021). A study of adolescents' and young adults' TikTok challenge participation in South India. *Human Factors in Healthcare*, 1, 100005. <https://doi.org/https://doi.org/10.1016/j.hfh.2022.100005>
- Rothstein, A., & Butler, C. T. (1987). *On Conflict and Consensus: A Handbook on Formal Consensus Decision Making*. Food Not Bombs.
- Sagi, O., & Rokach, L. (2021). Approximating XGBoost with an interpretable decision tree. *INFORMATION SCIENCES*, 572, 522-542. <https://doi.org/10.1016/j.ins.2021.05.055>

- Schneider, C. Q., & Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge University Press. <https://doi.org/DOI:10.1017/CBO9781139004244>
- Schwartz, S. H., & Sagie, G. (2000). Value Consensus and Importance. *Journal of Cross-Cultural Psychology*, 31(4), 465-497. <https://doi.org/10.1177/0022022100031004003>
- Stevens, S. S. (1966). A metric for the social consensus. *Science*, 151 3710, 530-541.
- Sunstein, C. R. (2002). The Law of Group Polarization. *Journal of Political Philosophy*, 10(2), 175-195. <https://doi.org/10.1111/1467-9760.00148>
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149-178. <https://doi.org/10.1002/ejsp.2420010202>
- Tran, T. N. T., Felfernig, A., & Le, V. M. (2024). An overview of consensus models for group decision-making and group recommender systems. *User Modeling and User-Adapted Interaction*, 34(3), 489-547. <https://doi.org/10.1007/s11257-023-09380-z>
- Wang, D., Kong, X., Nie, X., Shang, Y., Xu, S., He, Y., Maguire, P., & Hu, Y. (2021). The effects of emotion and social consensus on moral decision-making. *Ethics & Behavior*, 31(8), 575-588. <https://doi.org/10.1080/10508422.2020.1830404>
- Wu, T. (2024). Heterogeneous Opinion Dynamics Considering Consensus Evolution in Social Network Group Decision-Making. *Group Decision and Negotiation*, 33(1), 159-194. <https://doi.org/10.1007/s10726-023-09858-6>
- Zhang, M., & Du, Y. (2019). Qualitative Comparative Analysis (QCA) in Management and Organization Research: Position, Tactics, and Directions. *Chinese Journal of Management*, 16(09), 1312-1323.
- Zhang, Y., Chen, X., Gao, L., Dong, Y., & Pedrycz, W. (2022). Consensus reaching with trust evolution in social network group decision making. *Expert Systems with Applications*, 188, 116022. <https://doi.org/10.1016/j.eswa.2021.116022>
- Zhou, Q., Wu, Z., Altalhi, A. H., & Herrera, F. (2020). A two-step communication opinion dynamics model with self-persistence and influence index for social networks based on the DeGroot model. *INFORMATION SCIENCES*, 519, 363-381. <https://doi.org/10.1016/j.ins.2020.01.052>

© [CC-BY-NC 4.0](#) The Author(s). For more information, see our [Open Access Policy](#).