COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 50 (2026) e70252

ISSN: 1551-6709 online
DOI: 10.1111/cogs.70252

When Others Agree More Closely: Judgment Similarity Influences Conformity

Haokui Xu,[a] Xingyuan Zhang,[b,c] Jiaqi Wu,[b,c] Jun Yin[b,c]

[a]*Department of Psychology, College of Education, Zhejiang University of Technology*
[b]*Department of Psychology, Ningbo University*
[c]*Center of Group Behavior and Social Psychological Service, Ningbo University*

Received 23 December 2025; received in revised form 8 July 2026; accepted 15 July 2026

Abstract

Conformity is strongly shaped by perceived consensus, yet prior research has predominantly examined this process in discrete-choice contexts, where consensus is typically operationalized as the size of a uniform majority. In continuous real-world environments, however, social judgments naturally vary, making judgment similarity a potentially important cue for consensus. The present study investigated whether similarity among others' judgments modulates conformity. Across two perceptual estimation experiments, participants first estimated the location of a briefly presented target and then revised their estimates after viewing two partners' responses. In Experiment 1, in which partner responses were attributed to humans, participants showed greater conformity when partners' estimates were highly similar. Computational modeling formalized this process within a Bayesian belief-updating framework, suggesting that judgment similarity increased the evidential weight assigned to social information. In Experiment 2, this similarity-dependent effect was attenuated when the same response structure was explicitly attributed to computer programs rather than human agents. These findings indicate that conformity in continuous judgment spaces cannot be fully explained by a domain-general statistical heuristic or passive imitation alone. Instead, individuals appear to evaluate both the content of social evidence and the attributed properties of its source, using judgment similarity as a cue for integrating information from others.

Keywords: Conformity; Judgment similarity; Social influence; Consensus

Correspondence should be sent to Jun Yin, Department of Psychology, Ningbo University, No. 616 Fenghua Rd., Nongbo 315211, China. E-mail: yinjun1@nbu.edu.cn

1. Introduction

When faced with perceptual or cognitive uncertainty, individuals frequently rely on social information from multiple others to update their beliefs and judgments—a fundamental process classically defined as conformity (Cialdini & Goldstein, 2004; Sherif, 1936). Although decades of research have primarily examined conformity in discrete choices contexts (Bond, 2005; Cialdini, Reno, & Kallgren, 1990; Latane, 1981; Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007), real-world social environments rarely involve perfectly uniform judgments. Instead, social partners often exhibit actions that vary subtly, reflecting different degrees of similarity. This raises the following important question: how does the similarity among social partners' judgments modulate conformity?

In the classic research on conformity, the strength of group consensus is considered a primary determinant (Asch, 1956; Kim & Hommel, 2015; Moussaïd et al., 2016; Park, Goïame, O'Connor, & Dreher, 2017; Zhang & Gläscher, 2020). Consensus refers to the degree of agreement among group members regarding a particular judgment and is typically operationalized as the absolute number or relative proportion of individuals endorsing the same discrete option (Bond, 2005; Latane, 1981; Morgan, Rendell, Ehn, Hoppitt, & Laland, 2012; Moussaïd et al., 2016; Park et al., 2017; Schultz et al., 2007). The influence of consensus has been demonstrated across diverse experimental paradigms and appears to be deeply embedded in human behavior. Even in anonymous, private judgment contexts, individuals tend to adjust their responses toward the group consensus (Deutsch & Gerard, 1955). This sensitivity to consensus also emerges early in development, as young children show majority-biased transmission (Haun, Rekers, & Tomasello, 2012). The effect of consensus on conformity is commonly attributed to both normative and informational mechanisms. From a normative perspective, high consensus can increase perceived social pressure because dissent from a majority may threaten social acceptance or group belonging (Cialdini & Goldstein, 2004). From an informational perspective, high consensus can serve as evidence for the validity or adaptiveness of a response, particularly when converging judgments are perceived as independent (Pfänder, De Courson, & Mercier, 2025). Together, these findings indicate that consensus level plays a critical role in shaping social influence.

However, applying a numerical index of consensus to continuous judgment domains is challenging. In continuous decision-making contexts, such as spatial-location estimation, independent judgments naturally vary, making perfect agreement unlikely. Although prior studies have examined conformity using continuous responses, such as facial attractiveness ratings or target position estimates (Klucharev, Hytönen, Rijpkema, Smidts, & Fernández, 2009; Mahmoodi, Bahrami, & Mehring, 2018), they have typically treated the average value as the reference point for conformity while giving less attention to the degree of similarity among individual judgments. In real-world settings, judgments are often continuous rather than categorical. For example, eyewitnesses may provide slightly different estimates of when an event occurred; investors may generate closely aligned or widely divergent forecasts of future price levels. In such contexts, higher judgment similarity reflects tighter convergence among informants, whereas lower judgment similarity reflects weaker convergence. Consequently, judgment similarity—defined as the spatial proximity or convergence among multiple continuous judgments—may serve as a fine-grained cue for assessing consensus and the reliability of social information. This perspective extends classical accounts of consensus in conformity research (Asch, 1956; Bond, 2005) to continuous spaces of social information. It also aligns with recent arguments that decision-making research should move beyond discrete-choice paradigms to better capture the continuous structure of real-world judgments (Rasanan et al., 2024). Examining how gradations in judgment similarity influence conformity is, therefore, critical for understanding the nuanced dynamics of social influence in complex social environments (Cikara, Martinez, & Lewis, 2022).

Seemingly, judgment similarity might appear to be unrelated to conformity because individuals often automatically imitate behaviors observed from multiple partners (Cracco & Cooper, 2019), with the average strength of these imitative responses potentially guiding whether they follow others' actions. However, humans also possess mechanisms of epistemic vigilance that allow them to evaluate communicated information rather than relying on rigid or passive imitation alone (Mercier & Miton, 2019; Orticio, Martí, & Kidd, 2022). From an informational influence perspective, consensus among others' choices can serve as a reliable cue to the potential validity of social information (Gardikiotis, Martin, & Hewstone, 2005; Morgan et al., 2012). Importantly, greater convergence among partners' judgments reflects stronger consensus (Pfänder et al., 2025). Observers may, therefore, treat highly convergent judgments as more credible and rely on them more strongly when forming their own judgments (Cracco et al., 2022). On the basis of this reasoning, we hypothesize that judgment similarity dynamically modulates conformity: individuals will show greater conformity when others' judgments are highly similar than when they are less similar.

To examine whether higher similarity among partners' judgments increases conformity, the present study adapted the paradigm introduced by Mahmoodi et al. (2018). Participants were asked to estimate the location of briefly flashing dots positioned on a circular ring. After their initial response, they were shown the responses of two partners and given the opportunity to revise their answers. The angular adjustment from the first to the second estimate served as the index of conformity. Experiment 1 examined whether judgment similarity among ostensible human partners increased conformity. Partners' responses were generated by a computer program and were configured to appear either spatially close to one another (high similarity) or spatially distant from one another (low similarity) on the response ring. At the same time, a credible human-social context was established to lead participants to believe that they were performing the task with real human partners. This design allowed us to isolate the effect of judgment similarity while preserving perceived social presence. Because informational social influence may depend not only on the content of social information but also on the source to which that information is attributed, Experiment 2 examined whether the effect of judgment similarity depends on the information source, specifically, human-generated information versus non-human-generated information. Partner responses were generated using the same program and structure as in Experiment 1, but participants were explicitly informed that these responses were produced by computer programs. This manipulation allowed us to isolate the contribution of perceived information source by comparing experiments with matched informational content but different source attributions. We further developed computational models to examine how judgment similarity contributes to informational conformity. A

Bayesian belief-updating framework is well suited for this purpose because it formalizes how beliefs and responses are updated on the basis of observed social evidence and has been widely used in perceptual decision-making and consensus research (Duan, Guo, Zheng, & Yin, 2026; Toelch & Dolan, 2015; Xie & Hayes, 2022). By representing latent cognitive quantities as model parameters, this framework allowed us to simulate alternative psychological strategies and evaluate which strategy better accounted for participants' response patterns. In the present study, we implemented multiple models within the same Bayesian framework, incorporating individual confidence, self-other difference, and judgment similarity in different combinations. Parameters were estimated using maximum likelihood estimation, and models were compared using the Bayesian information criterion (BIC) to test whether including judgment similarity improved explanatory performance after penalizing model complexity.

2. Experiment 1

We adapted a perceptual task based on Mahmoodi et al. (2018) to investigate the effect of judgment similarity on conformity. Participants estimated the location of a visual target on a computer screen, and then revised their estimate after observing two partners' estimates for the same target. Critically, the two partner responses were generated by a computer program. On some trials, the partners' estimates were spatially close to one another, corresponding to the high-similarity condition; on other trials, they were farther apart, corresponding to the low-similarity condition. Participants, however, were not informed that the partner responses were computer-generated. Instead, they were led to believe that they were completing the task with real human partners.

2.1. Methods

2.1.1. Transparency and openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. Although both experiments were not preregistered, the sample size determination, exclusion criteria, and an initial analysis plan were specified before data collection. Sample size was determined a priori using G*Power (version 3.1, Faul, Erdfelder, Buchner, & Lang, 2009). Behavioral analyses were conducted in SPSS Statistics (version 25.0.0.0) and R, and computational modeling was implemented in Python. The initial analysis plan included (a) a *t*-test for the focal effect of judgment similarity(b) repeated-measures ANOVA for interaction tests involving similarity and other task-related factors, and (c) linear mixed-effects (LME) models for confidence-related effects, given the imbalance in trial counts across confidence ratings. To provide a more consistent and statistically robust inferential framework, and to reduce ambiguity arising from the use of multiple analytic approaches, we subsequently harmonized the primary analyses using LME models. We retained the planned *t*-test for the core similarity effect because the a priori power analysis was based on this test. Detailed descriptions of sample-size calculation, exclusion criteria, the initial analysis plan, and the revised analysis framework are provided in the sections below.

2.1.2. *Participants*

The sample size was determined via a G*Power analysis (version 3.1, Faul et al., 2009) for detecting a conventional medium effect size (*Cohen's* $d = 0.5$) with a paired *t*-test, an alpha level of 0.05, and a power of 0.80; the required sample size was found to be at least 34 participants. Finally, 36 campus participants (19 males, 17 females; $M_{\text{age}} = 21.11$ years, $SD_{\text{age}} = 1.54$) took part in Experiment 1 in exchange for 45 RMB. We maintained a consistent sample size across all the experiments. The a priori power analysis targeted the core similarity contrast and was not intended to guarantee sensitivity for all interaction effects.

2.1.3. *Experimental procedures and design*

Participants sat approximately 60 cm from a 15.6-inch monitor (2560 × 1600 resolution, 90 Hz refresh rate). At the beginning of the first trial, they first viewed a rapid sequence of 91 Gaussian dots (radius = 5 mm): the first stimulus was presented for 30 ms, while every other stimulus was presented for 15 ms each. Once all stimuli disappeared, participants indicated the location of the first dot by clicking on a circular ring using a mouse. This first estimate was highlighted in a yellow circle. They then rated their confidence on a 1–6 scale (1 = low confidence, 6 = high confidence). Next, the estimates of two partners were displayed simultaneously (highlighted in red circles), after which participants had the opportunity to revise their initial response (second response, highlighted as a yellow square). The trial ended after the second click, followed by a 1–1.5 s inter-trial interval. No time limits were imposed on participants' responses (see Fig. 1a).

The experiment used a 2 × 3 within-subject design (60 trials total). The first factor was partners' judgment similarity, which was manipulated via the angular distance between the two partner estimates. This factor included two levels: high similarity, in which partner estimates were 24°–36° apart, and low similarity, in which partner estimates were 54°–66° apart. Each similarity level included 30 trials. The second factor was self-other difference, defined as the angular distance between the participant's first estimate and the mean of the two partner estimates. This factor included three levels: 60°, 90°, and 120°, with 10 trials per level within each similarity condition. Trial order was randomized for each participant.

Critically, partner estimates were generated programmatically in accordance with the design constraints, although participants in Experiment 1 were not informed of this procedure. To maintain the belief that they were interacting with real human partners, three participants arrived at the laboratory simultaneously and were assigned to separate computer-equipped compartments in the same room. The compartments were divided by partitions, preventing participants from seeing one another, and participants were instructed not to communicate during the task (the full room layout and participant-separation setup are provided in Fig. S1). An experimenter then entered each compartment to provide the same instructions. Participants were explicitly informed that partners' responses would be displayed only after all first-round estimates had been submitted. Thus, from the participants' perspective, there was no reason to infer that one partner's initial estimate had been influenced by the other partner. Participants were also informed that all players would provide a second-round response; however, these second-round responses would not be visible to the other players. Thus, participants understood that their own revisions would not be observed by their partners.

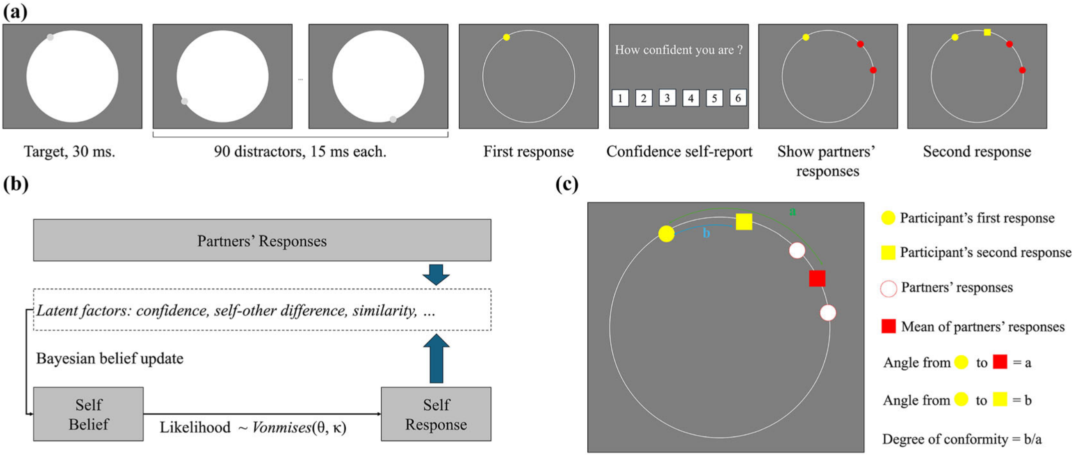

Fig. 1. (a). Experimental task. Participants observed a sequence of dots on the screen and were asked to estimate the location of the first dot, indicated by a yellow circle. They then reported their confidence on a scale from 1 (low) to 6 (high). Next, the estimates of two partners for the same stimulus were shown, marked by red circles. After viewing these partner responses, participants could revise their estimate, shown as a yellow square. (b). Computational model. A unified Bayesian belief-updating framework was used to model how participants revise their estimates after observing social information. Candidate models differed in the cues that modulated the weighting of social evidence, including confidence, self-other difference, judgment similarity, and nearest-neighbor-based cues. Model parameters were estimated using maximum likelihood estimation, and model fit was compared using the Bayesian information criterion (BIC). (c). Measurement of conformity. Conformity was quantified by the angular change between the participant's first and second responses, divided by the angular distance between the participant's initial response and the mean of the partners' responses (red square).

Together, these procedures were designed to preserve a credible social-interaction context while minimizing perceived dependency between partners' estimates and audience-related concerns, such as politeness or reputation management.

2.1.4. Analysis

We quantified the conformity by calculating the extent to which participants adjusted their responses toward their partners'. Specifically, the conformity was measured as the angular difference between a participant's first and second clicks, divided by the difference between the participant's and the average of the partner's initial responses (Fig. 1c). A higher value indicated a stronger degree of conformity.

In the initial analysis plan, we used a paired-samples *t*-test for the focal judgment-similarity contrast, and ANOVA-based analyses for condition-level interaction checks. Because confidence-related trial counts were imbalanced, we additionally specified mixed-effects analyses for confidence-involved effects. More details and results from the initial analysis plan are reported in Section S4 (Tables S9 and S10).

In the final analysis plan, we retained the paired-samples *t*-test for the core similarity contrast as a planned baseline analysis linked to the a priori power calculation. The primary inferential analyses were conducted using LME models, with conformity as the dependent variable, self-other difference and confidence as continuous predictors. Judgment similarity was coded using orthogonal contrast coding (low $= -0.5$, high $= 0.5$). Both self-other difference and confidence were mean-centered. The random effect included random intercepts of participants. To test whether the effects generalized across human source and nonhuman source contexts, the cross-experiment model additionally included experiment (Exp1 vs. Exp2; sum-coded) and all interactions among experiment, similarity, self-other difference, and confidence. Fixed effects were evaluated using Type III ANOVA with Satterthwaite approximations for denominator degrees of freedom. For significant or theoretically relevant interactions, we conducted planned simple-effects and simple-slopes analyses using estimated marginal means at centered values and at ± 1 SD probe points for continuous moderators.

2.1.5. *Computational modeling*

We implemented two computational models grounded in a Bayesian belief-updating framework to simulate the cognitive mechanisms underlying conformity. In this framework, participants were assumed to hold an initial prior belief about the target location and to update this belief after observing social evidence, operationalized as the partners' responses (Eq. 1). The revised estimate was modeled as a sample from a von Mises distribution centered on the resulting posterior belief.

$$\hat{\theta}_t^{posterior} \sim (1 - w_t)\theta_t^{self} + w_t \theta_t^{social} \tag{1}$$

where $\hat{\theta}_t^{posterior}$ indicates the model-predicted posterior on trial t; θ_t^{self} indicates the initial judgment of the participants on trial t; θ_t^{social} indicates the social estimate (from partners' judgments) on trial t; and w_t refers to the weighting parameter of social evidence, bounded between 0 and 1, indicating how strongly social information influences updating.

To test our hypothesis that judgment similarity modulates conformity, we compared two nested models in which social evidence is weighted in different ways: Basic model (BM). The BM assumed that social information was represented by the mean of the two partners' estimates. The weight assigned to this social evidence was specified as a function of two cues: the participant's self-reported confidence and the angular difference between their initial estimate and the partners' mean estimate, referred to as self-other difference (Eq. 2).

$$w_t \sim \beta_1 Z_{self-other,t} + \beta_2 Z_{confidence,t} \tag{2}$$

Similarity-weighted model (SWM). The SWM extended the BM by incorporating judgment similarity between the two partners' estimates as an additional cue. This model assumed that the evidential weight assigned to the social mean was dynamically adjusted as a function of the degree of convergence between partners' estimates (Eq. 3).

$$w_t \sim \beta_1 Z_{self-other,t} + \beta_2 Z_{confidence,t} + \beta_3 Z_{similarity,t} \tag{3}$$

Free parameters for both models were estimated at the pooled level using maximum likelihood estimation, implemented by minimizing the negative log-likelihood (NLL) of participants' revised estimates. To evaluate whether adding the similarity-related parameter improved model fit after accounting for model complexity, we compared models using the

Table 1
Fixed-effects coefficient estimates (LME) in Experiment 1

Effect/parameter	Estimate (β)	95% CI	t	p
Intercept	0.364	[0.312, 0.417]	13.62	< .001
Judgment similarity(high = 0.5, low = −0.5)	−0.141	[−0.168, −0.115]	−10.54	< .001
Self-other difference	0.002	[0.001, 0.002]	6.79	< .001
Confidence	−0.112	[−0.125, −0.099]	−16.89	< .001
Judgment similarity × Self-other difference	−0.001	[−0.002, 0.000]	−1.84	.065
Judgment similarity × Confidence	0.034	[0.015, 0.054]	3.49	< .001
Self-other difference × Confidence	−0.001	[−0.001, −0.000]	−2.01	.044
Judgment similarity × Self-other difference × Confidence	0.001	[−0.000, 0.001]	1.24	.215

BIC. The model with the lower BIC was interpreted as providing the better fit to the empirical data. Likelihood specification and model-comparison equations are provided in Section S5.

2.2. Results

2.2.1. Behavioral results

Data from 36 participants were included in the final analysis. Two additional participants also took part in the experiment but were not included in the analysis because they expressed doubts about the authenticity of the partners' estimates. When the judgment similarity between partners was high, participants were more likely to conform ($M = 0.43$; $SE = 0.04$) than when the judgment similarity was low ($M = 0.29$, $SE = 0.03$; $t(35) = 6.02$, $p < .001$, *Cohen's d* = 1.00, 95% CI of *Cohen's d* = [0.60, 1.40]).

Trial-level LME analyses converged on this pattern: judgment similarity significantly predicted conformity ($F(1, 2122.89) = 111.08$, $p < .001$, $\eta_p{}^2 = 0.05$). The negative coefficient for judgment similarity ($b = -0.141$, $t = -10.54$, $p < .001$) indicated that participants showed greater conformity when partners' judgments were more similar. Beyond this core effect, self-other difference and confidence were also significant predictors, and the interaction between judgment similarity and confidence was significant, indicating that the similarity effect varied as a function of confidence. We did not detect reliable evidence for the three-way interaction ($F(1, 2119.85) = 1.54$, $p = .215$, $\eta_p{}^2 = 0.001$). However, higher-order interactions were not reliably detected and should be interpreted cautiously (sensitivity analysis for higher-order interaction terms could be found in Table S8). Full fixed-effect estimates and 95% CIs are reported in Table 1, and planned simple-effects/slopes analyses are reported in the Supplementary Materials (Tables S1–S3).

2.2.2. Modeling results

The SWM showed a lower NLL (1975.85) and lower BIC (3982.40) than the BM (NLL = 2042.36, BIC = 4107.75). Therefore, even after the complexity penalty, the SWM provided a better fit to the empirical data than the BM (ΔBIC = 125.35). This result supported the theoretical prediction that judgment similarity modulated conformity.

To clarify the mechanisms captured by the best-fitting SWM, we examined its maximum likelihood parameter estimates. The model estimated separate weighting coefficients for self-other difference, confidence, and judgment similarity. The weighting coefficients for the self-other difference (β_1), confidence (β_2), and partner judgment similarity (β_3) were estimated at 0.141, 0.776, and 0.435, respectively. Crucially, the positive parameter value for judgment similarity (β_3= 0.435) indicates that greater similarity between partners' estimates increased the evidential weight assigned to social information, thereby producing stronger conformity.

Grounded in our theory-driven hypotheses, the main-text model comparison focused on BM versus SWM as the primary test of whether judgment similarity influences conformity. To seek a more complete account of the underlying psychological process and improve overall fit, we additionally fit a broader model set that included a nearest-neighbor strategy (see Section S5 and Table S11). In this expanded comparison, the nearest-neighbor model showed lower BIC (3698.67) than SWM, indicating that nearest-neighbor heuristics capture additional variance in participants' trial-level adjustments. Importantly, adding judgment similarity further reduced BIC (with the similarity-weighted nearest neighbor model achieving the lowest BIC overall, BIC = 3659.80), which reinforces the importance of judgment similarity.

To evaluate absolute fit quality, we computed observed-versus-predicted fit indices and visualization diagnostics (see Section S5 and Fig. S2). These results indicate that the better-performing models capture systematic variation in behavior at a moderate level. Consistent with our core claim, models that include judgment similarity tend to provide better performance than their corresponding counterparts without judgment similarity, indicating that judgment similarity contributes useful explanatory information in the updating process. However, even the better models retain nontrivial trial-level error, suggesting that the current model family does not provide a high-precision reconstruction of participants' responses.

2.3. Discussion

Experiment 1 demonstrated that judgment similarity modulates conformity: individuals were more likely to revise their estimates when their partners' estimates showed high similarity. Computational modeling further indicated that similarity-dependent conformity reflects a process of weighted belief updating, whereby greater judgment similarity increases the evidential weight assigned to partners' judgments.

3. Experiment 2

Experiment 1 showed that greater judgment similarity between partners increased the magnitude of participants' conformity. Computational modeling further suggested that judgment similarity modulated the weight assigned to observed social information during belief updating. However, belief updating may depend not only on the content of the information itself but also on the perceived properties of its source. On the basis of Experiment 1 alone, the mechanism underlying the judgment similarity effect remains unclear. Specifically, the effect may reflect a domain-general response to informational content, such that individuals treat highly similar judgments as more reliable evidence regardless of source. Alternatively, the effect may depend on social source attribution, such that judgment similarity increases conformity primarily when the information is believed to originate from human partners.

To investigate whether the effect of judgment similarity depends on the information source, specifically, human-generated information versus non-human-generated information, Experiment 2 used the same task structure and variable manipulations as Experiment 1, except that participants were explicitly informed that the partner estimates were generated by computer programs. This design held the informational content constant while manipulating the attributed source of the information. If the effect of judgment similarity depends solely on informational content, a comparable similarity effect should be observed in Experiment 2. Conversely, if the interaction between experiment (Experiment 1 vs. Experiment 2) and judgment similarity is significant, this would indicate that the similarity effect depends on source attribution.

3.1. Methods

Another 36 campus participants (21 males, 15 females; $M_{\text{age}} = 20.78$ years, $SD_{\text{age}} = 1.81$) took part in Experiment 2 in exchange for 45 RMB.

The experimental procedure closely mirrored that of Experiment 1, with the following modifications. Participants were informed that the partners' estimates were generated by computer algorithms, and the delay at the beginning of each trial used in Experiment 1—to reinforce their belief that the partners were human—was removed.

3.2. Results

3.2.1. Behavioral results

As shown in Fig. 2b, the conformity of the participants did not differ significantly between the high ($M = 0.54$, $SE = 0.05$) and low ($M = 0.51$, $SE = 0.05$) similarity conditions (in *t*-test: $t(35) = 1.63$, $p = .112$, *Cohen's d* $= 0.27$, 95% CI of *Cohen's d* $= [-0.06, 0.60]$; in mixed-effects linear model: $F(1, 2118.25) = 3.79$, $p = .052$, $\eta_p^2 = 0.002$). Confidence remained a significant predictor of conformity, whereas self-other difference and interaction terms were not significant. Full fixed-effect estimates and 95% CIs are reported in Table 2. Complete fixed-effect outputs for Experiment 2 are provided in Tables S4 and S5.

3.2.2. Between-experiment comparison

To test whether the similarity effect was specific to information attributed to human partners, we fitted a trial-level LME model to the pooled data from both experiments, with experiment (Exp1 vs. Exp2), judgment similarity, self-other difference, and confidence entered as fixed effects (including interactions), and a random intercept for participant nested within experiment. Type III tests showed significant main effects of experiment ($F(1, 69.78) = 5.79$, $p = .019$, $\eta_p^2 = 0.08$), and judgment similarity ($F(1, 4239.10) = 74.98$, $p < .001$, $\eta_p^2 = 0.02$). Critically, the experiment $\times$ similarity interaction was significant ($F(1, 4239.10) = 36.05$, $p < .001$, $\eta_p^2 = 0.01$), indicating that the similarity effect differed across experiments. The significant similarity effect was found in Experiment 1 ($p < .001$), but not in Experiment 2

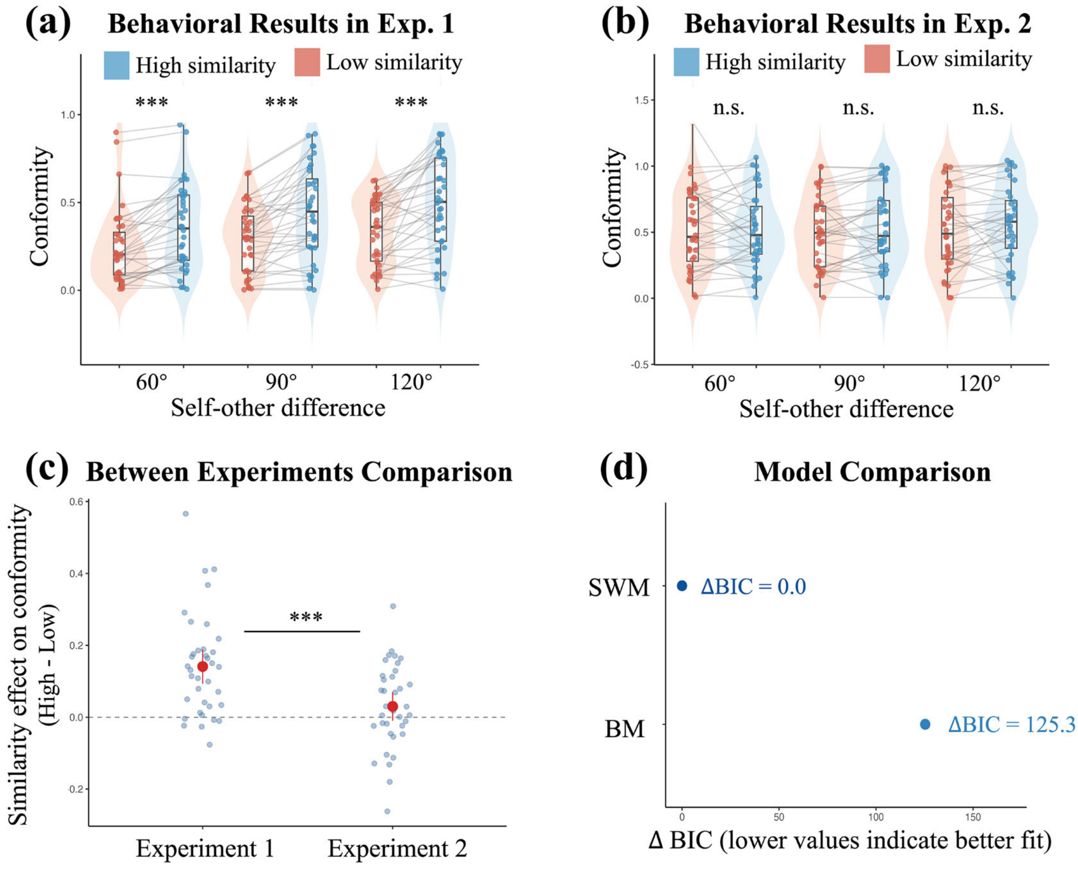

Fig. 2. Behavioral and computational modeling results across experiments. Subplots (a) and (b) present behavioral results from Experiments 1 and 2, respectively. Conformity is plotted as a function of self-other difference condition (60°, 90°, or 120°) and partners' judgment similarity (low vs. high). Violin plots show distribution across participants, boxplots indicate the median and interquartile range, points represent participant-level means, and gray lines connect within-participant values for the low and high similarity conditions within each self-other difference condition. (c) Shows the cross-experiment comparison of the similarity effect on conformity, computed for each participant as conformity in the high-similarity condition minus conformity in the low-similarity condition and then compared between Experiments 1 and 2. Points represent individual participants; red points and error bars indicate group means and 95% confidence intervals. (d) Shows model comparison between the basic model (BM) and the similarity-weighted model (SWM) using Δ BIC values computed relative to the best-fitting model. Lower Δ BIC indicates better fit after accounting for model complexity. In all subplots, *** indicates $p < .001$ and n.s. indicates a nonsignificant difference.

($p = .060$). This pattern suggests that judgment similarity influenced conformity primarily when partner estimates were attributed to human sources. Full fixed-effect estimates and 95% CIs are reported in Table 3. Full Type III tests and random-effects components remain in Tables S6 and S7.

Table 2
Fixed-effects coefficient estimates (LME) in Experiment 2

Effect/parameter	Estimate (β)	95% CI	t	p
Intercept	0.523	[0.433, 0.613]	11.38	< .001
Judgment similarity	−0.029	[−0.057, 0.000]	−1.95	.052
Self-other difference	0.001	[−0.000, 0.001]	1.22	.222
Confidence	−0.088	[−0.103, −0.073]	−11.50	< .001
Judgment similarity × Self-other difference	−0.001	[−0.002, 0.001]	−0.85	.397
Judgment similarity × Confidence	0.014	[−0.009, 0.036]	1.18	.237
Self-other difference × Confidence	−0.001	[−0.001, 0.000]	−0.44	.661
Judgment similarity × Self-other difference × Confidence	−0.001	[−0.001, 0.000]	−0.98	.329

Table 3
Trial-level LME fixed-effects estimates for cross-experiment comparison

Effect/parameter	Estimate (β)	95% CI	t	p
Intercept	0.446	[0.394, 0.498]	16.76	< .001
Experiment (Exp1 = 0.5, Exp2 = −0.5)	0.128	[0.024, 0.232]	2.41	.019
Judgment similarity	−0.087	[−0.106, −0.067]	−8.66	< .001
Experiment × Similarity	0.120	[0.081, 0.159]	6.00	< .001

3.3. Discussion

Experiment 2 highlighted the role of source attribution in the effect of judgment similarity on conformity. When the same response patterns of partners that had been attributed to human partners in Experiment 1 were instead described as computer-generated, the effect of judgment similarity on conformity was no longer observed. These findings suggest that conformity is shaped not only by the degree of similarity among others' responses but also by how the source of those responses is perceived.

4. General discussion

The present study extends classic research on consensus and conformity by introducing an underexamined factor in continuous social judgment: the similarity among others' judgments. Across two experiments, participants exhibited greater conformity when ostensible human partners provided highly similar estimates (Experiment 1). Crucially, this similarity-dependent effect was attenuated when the same response structure was described as computer-generated rather than human-generated (Experiment 2). Computational modeling formalized this process and suggested that similarity-dependent conformity is consistent with a weighted belief-updating account, in which greater judgment similarity increases the weight assigned to social evidence. Together, these findings indicate that conformity in continuous judgment contexts is shaped not only by the content of others' responses but also by their similarity and perceived source, highlighting an active evaluation of social information rather than passive imitation.

Conformity has traditionally been examined in terms of consensus strength in discrete-choice contexts, in which larger majorities increase the likelihood of alignment with the group (Asch, 1956). Robust majority-size effects (Bond, 2005; Latane, 1981) and relative-majority effects (Morgan et al., 2012; Tanford & Penrod, 1984) have also motivated a more restrictive definition of conformity in the social-learning literature: individuals are considered conformist when they follow the majority disproportionately, that is, more than expected from the observed frequency of others' judgment (Kendal et al., 2018). In the continuous space examined here, however, this strict operationalization is difficult to apply because it is not straightforward to define disproportionate copying for graded responses. We, therefore, adopt a broader, traditional definition of conformity as directional adjustment toward the judgments of multiple others (Cialdini & Goldstein, 2004; Cialdini & Trost, 1998). Under this definition, the present findings suggest that consensus enhancement induced by judgment similarity can increase conformity, extending conformity research to a more ecologically realistic continuous-response context. In real-world contexts, individuals rarely produce perfectly identical judgments; rather, their responses typically vary by degree. Nevertheless, stronger group consensus is often associated with greater response similarity (Allen & Levine, 1969; Hawkins, Goodman, & Goldstone, 2019; Kenny & La Voie, 1985). Observers may, therefore, treat high judgment similarity as a cue to stronger consensus and, accordingly, rely more heavily on social information.

From the perspective of informational social influence, the effect of judgment similarity on conformity reflects how humans adapt to uncertainty by integrating social information. Others' responses serve as informational cues that can help individuals reduce uncertainty and improve decision-making (Cohen & Golden, 1972; Mahmoodi, Nili, Bang, Mehring, & Bahrami, 2022). Crucially, these cues are not used uncritically but are evaluated for their reliability before relying on them (Mercier & Miton, 2019; Mercier & Morin, 2019; Sperber et al., 2010). The present results indicate that participants treated higher judgment similarity as a more reliable signal and, accordingly, showed greater conformity. This interpretation is consistent with recent work showing that, under broad conditions, more convergent responses across multiple informants are more likely to be accurate and to reflect greater competence (Pfänder et al., 2025). To further specify how this process operates within an informational social influence framework, we used computational modeling based on a unified Bayesian belief-updating architecture. Model comparison showed that adding judgment similarity improved model fit beyond confidence and self-other difference, indicating its incremental explanatory value. We also examined alternative strategies, including nearest-neighbor heuristics, in which individuals adjust toward the partner whose estimate is closest to their own initial response. Accordingly, we interpret BM/SWM as hypothesis-driven baseline mechanisms, rather than as an exclusive account of behavior, while the nearest-neighbor advantage identifies an important direction for mechanism refinement in future work. Importantly, even within these alternative model families, incorporating judgment similarity improved model fit (see Section S5, Table S10, and Fig. S2 for full model-space comparisons and fit diagnostics). Taken together, these findings suggest that judgment similarity serves as a useful cue to evidential reliability, shaping how individuals weight social information when updating their beliefs. It is worth noting that model fit remains only moderate and leaves substantial room for improvement (Fig. S2), suggesting that additional mechanisms not captured by the current model set likely contribute to trial-level behavior and warrant further investigation.

Within an informational social influence framework, evaluations of reliability are shaped not only by informational content but also by the attributed source of that information (Sperber et al., 2010; Yin et al., 2024). The cross-experiment results showed that the effect of judgment similarity was attenuated when participants believed that partner responses were generated by computer programs rather than human agents. This pattern suggests that source attribution modulates the extent to which similarity is used during belief updating. One plausible account is that participants use source cues to infer how information was generated. When responses were attributed to human agents, high similarity may have been interpreted as evidence of independent convergence or shared competence and, therefore, assigned greater evidential weight (Pfänder et al., 2025). By contrast, when the same response pattern is attributed to computer generation, participants may have been less likely to treat sources as epistemically independent with respect to competence, even if the response-generation process is technically independent. Instead, they may have inferred that the responses originated from the same underlying program logic and interpreted the observed variation as systematic or random error that could be reduced by averaging. This inference may have reduced the diagnostic value of judgment similarity for reliability assessment. It is important to clarify that, in Experiment 2, participants were told only that partner responses were generated by computer programs; the source was not described as artificial intelligence. Participants may, therefore, have been more likely to infer conventional computer programs rather than contemporary large-scale artificial intelligence (AI) systems. Given growing evidence that advanced AI can elicit human-like competence attributions and anthropomorphic representations, the present data cannot determine whether similarity would be used in the same way when responses are attributed to AI agents. The appropriate conclusion from the present evidence is that source attribution shapes how individuals use judgment similarity to evaluate the reliability of social information.

Although our theoretical account and computational modeling primarily follow an informational social influence framework, we do not rule out the possibility that normative processes also contributed to the observed effects. Judgment similarity may shape perceptions of normative expectations within a group: when partners provide highly similar responses, their convergence may be interpreted as evidence of a stronger norm of consistency, thereby increasing conformity (Hawkins et al., 2019; Wang & Liao, 2023). This account may also help explain why the effect of judgment similarity was attenuated when partners were described as computer programs. Because computers are less likely to be perceived as members of a relevant social group, their responses may carry less normative weight, reducing the influence of judgment similarity on conformity (Kalish, 2012). However, because the present perceptual task did not directly measure perceived norms, this interpretation remains tentative and should be examined in future research.

Several limitations of the present study should be acknowledged. First, the number of partners was fixed at two, leaving open whether judgment similarity would exert comparable effects when larger groups are observed. Second, in Experiment 2, participants were told that the responses were computer-generated, but we did not examine potential moderators such as algorithmic trust or perceived algorithmic reliability. Relatedly, we did not directly assess participants' impressions of the partners, such as perceived competence or trustworthiness. Although interpersonal exposure was minimized, the influence of such impressions cannot be fully ruled out. Third, although procedures were designed to ensure that participants in Experiment 1 believed their partners were real, the partners' responses were in fact computer-generated, and the task did not include performance-contingent incentives. Thus, the impact of dynamic, real-time interactions among human partners, and stronger accuracy motivation on conformity was not assessed, which may limit the ecological validity of the findings. Future work should test larger groups, directly measure source-related impressions and trust, and use more naturalistic interactive designs with incentive-compatible performance settings.

Taken together, the present study reveals that conformity is shaped not only by the number of individuals endorsing a given response but also by the degree of similarity among their judgments. Participants showed greater conformity when partners' judgments were highly similar, but this effect was observed primarily when those judgments were attributed to human rather than computer-generated sources. These findings suggest that conformity reflects evaluative social-information use rather than passive imitation alone: Individuals integrate not only the content of others' judgments but also cues to their reliability and source. Such integration may help individuals extract informational value from social environments and adapt their judgments under uncertainty.

Funding

This research was supported by the National Natural Science Foundation of China (grant no. 32371090) and the Zhejiang Provincial Natural Science Foundation of China (grant no. LQN26C090003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of interest

The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval statement

The study was approved by the Research Ethics Board of the Department of Psychology at Ningbo University. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent to participate and consent publication

Informed consent was obtained from all individual participants included in the study. Patients signed informed consent regarding publishing their data.

Availability of data, code, and materials

Data and code have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/q2zet/?view_only=f22a52b1127349c48655497f4b018d46.

References

Allen, V. L., & Levine, J. M. (1969). Consensus and conformity. *Journal of Experimental Social Psychology*, *5*(4), 389–399. https://doi.org/10.1016/0022-1031(69)90032-8

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*(9), 1–70. https://doi.org/10.1037/h0093718

Bond, R. (2005). Group size and conformity. *Group Processes and Intergroup Relations*, *8*(4), 331–354. https://doi.org/10.1177/1368430205056464

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*(1), 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Ed.), *The handbook of social psychology* (4th ed., Vol. *2*, pp. 151–192). McGraw-Hill.

Cikara, M., Martinez, J. E., & Lewis, N. A. (2022). Moving beyond social categories by incorporating context in social psychological theory. *Nature Reviews Psychology*, *1*(9), 537–549. https://doi.org/10.1038/s44159-022-00079-3

Cohen, J. B., & Golden, E. (1972). Informational social influence and product evaluation. *Journal of Applied Psychology*, *56*(1), 54–59. https://doi.org/10.1037/h0032139

Cracco, E., Bernardet, U., Sevenhant, R., Vandenhouwe, N., Copman, F., Durnez, W., Bombeke, K., & Brass, M. (2022). Evidence for a two-step model of social group influence. *iScience*, *25*(9), 104891. https://doi.org/10.1016/j.isci.2022.104891

Cracco, E., & Cooper, R. P. (2019). Automatic imitation of multiple agents: A computational model. *Cognitive Psychology*, *113*, 101224. https://doi.org/10.1016/j.cogpsych.2019.101224

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, *51*(3), 629–636. https://doi.org/10.1037/h0046408

Duan, J., Guo, X., Zheng, L., & Yin, J. (2026). A Bayesian perspective on observers' inference of group norms. *Npj Science of Learning*, *11*, 24. https://doi.org/10.1038/s41539-026-00405-x

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Gardikiotis, A., Martin, R., & Hewstone, M. (2005). Group consensus in social influence: Type of consensus information as a moderator of majority and minority influence. *Personality and Social Psychology Bulletin*, *31*(9), 1163–1174. https://doi.org/10.1177/0146167205277807

Haun, D. B. M., Rekers, Y., & Tomasello, M. (2012). Majority-biased transmission in chimpanzees and human children, but not orangutans. *Current Biology*, *22*(8), 727–731. https://doi.org/10.1016/j.cub.2012.03.006

Hawkins, R. X. D., Goodman, N. D., & Goldstone, R. L. (2019). The emergence of social norms and conventions. *Trends in Cognitive Sciences*, *23*(2), 158–169. https://doi.org/10.1016/j.tics.2018.11.003

Kalish, C. W. (2012). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology*, *48*(4), 1133–1143. https://doi.org/10.1037/a0026344

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. *Trends in Cognitive Sciences*, *22*(7), 651–665. https://doi.org/10.1016/j.tics.2018.04.003

Kenny, D. A., & La Voie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology*, *48*(2), 339–348. https://doi.org/10.1037/0022-3514.48.2.339

Kim, D., & Hommel, B. (2015). An event-based account of conformity. *Psychological Science*, *26*(4), 484–489. https://doi.org/10.1177/0956797614568319

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., & Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, *61*(1), 140–151. https://doi.org/10.1016/j.neuron.2008.11.027

Latane, B. (1981). The psychology of social impact. *American Psychologist*, *36*(4), 343–356.

Mahmoodi, A., Bahrami, B., & Mehring, C. (2018). Reciprocity of social influence. *Nature Communications*, *9*(1), 2474. https://doi.org/10.1038/s41467-018-04925-y

Mahmoodi, A., Nili, H., Bang, D., Mehring, C., & Bahrami, B. (2022). Distinct neurocomputational mechanisms support informational and socially normative conformity. *PLOS Biology*, *20*(3), e3001565. https://doi.org/10.1371/journal.pbio.3001565

Mercier, H., & Miton, H. (2019). Utilizing simple cues to informational dependency. *Evolution and Human Behavior*, *40*(3), 301–314. https://doi.org/10.1016/j.evolhumbehav.2019.01.001

Mercier, H., & Morin, O. (2019). Majority rules: How good are we at aggregating convergent opinions? *Evolutionary Human Sciences*, *1*, e6. https://doi.org/10.1017/ehs.2019.6

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1729), 653–662. https://doi.org/10.1098/rspb.2011.1172

Moussaïd, M., Kapadia, M., Thrash, T., Sumner, R. W., Gross, M., Helbing, D., & Hölscher, C. (2016). Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of the Royal Society Interface*, *13*(122), 20160414. https://doi.org/10.1098/rsif.2016.0414

Orticio, E., Martí, L., & Kidd, C. (2022). Social prevalence is rationally integrated in belief updating. *Open Mind*, *6*, 77–87. https://doi.org/10.1162/opmi_a_00056

Park, S. A., Goïame, S., O'Connor, D. A., & Dreher, J.-C. (2017). Integration of individual and social information for decision-making in groups of different sizes. *PLOS Biology*, *15*(6), e2001958. https://doi.org/10.1371/journal.pbio.2001958

Pfänder, J., De Courson, B., & Mercier, H. (2025). How wise is the crowd: Can we infer people are accurate and competent merely because they agree with each other? *Cognition*, *255*, 106005. https://doi.org/10.1016/j.cognition.2024.106005

Rasanan, A. H. H., Evans, N. J., Fontanesi, L., Manning, C., Huang-Pollock, C., Matzke, D., Heathcote, A., Rieskamp, J., Speekenbrink, M., Frank, M. J., Palminteri, S., Lucas, C. G., Busemeyer, J. R., Ratcliff, R., & Rad, J. A. (2024). Beyond discrete-choice options. *Trends in Cognitive Sciences*, *28*(9), 857–870. https://doi.org/10.1016/j.tics.2024.07.004

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434.

Sherif, M. (1936). *The psychology of social norms*. Harper & Row.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Tanford, S., & Penrod, S. (1984). Social Influence Model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin*, *95*(2), 189–225. https://doi.org/10.1037/0033-2909.95.2.189

Toelch, U., & Dolan, R. J. (2015). Informational and normative influences in conformity from a neurocomputational perspective. *Trends in Cognitive Sciences*, *19*(10), 579–589. https://doi.org/10.1016/j.tics.2015.07.007

Wang, E., & Liao, Y.-T. (2023). Effects of member similarity on group norm conformity, group identity and social participation in the context of social networking sites. *Internet Research*, *34*(3), 868–890. https://doi.org/10.1108/INTR-09-2021-0632

Xie, B., & Hayes, B. (2022). Sensitivity to evidential dependencies in judgments under uncertainty. *Cognitive Science*, *46*(5), e13144. https://doi.org/10.1111/cogs70252.13144

Yin, J., Xu, Z., Lin, J., Zhou, W., & Guo, X. (2024). Smartly following others: Majority influence depends on how the majority behavior is formed. *Journal of Experimental Social Psychology*, *115*, 104644. https://doi.org/10.1016/j.jesp.2024.104644

Zhang, L., & Gläscher, J. (2020). A brain network supporting social influences in human decision-making. *Science Advances*, *6*(34), eabb4159. https://doi.org/10.1126/sciadv.abb4159

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supplementary Information