

How truth wins in opinion dynamics along issue sequences

Noah E. Friedkin^{a,b,1} and Francesco Bullo^{a,c}

^aCenter for Control, Dynamical Systems and Computation, University of California, Santa Barbara, CA 93106; ^bDepartment of Sociology, University of California, Santa Barbara, CA 93106; and ^cDepartment of Mechanical Engineering, University of California, Santa Barbara, CA 93106

Edited by Matthew O. Jackson, Stanford University, Stanford, CA, and approved September 7, 2017 (received for review June 13, 2017)

How truth wins in social groups is an important open problem. Classic experiments on social groups dealing with truth statement issues present mixed findings on the conditions of truth abandonment and reaching a consensus on the truth. No theory has been developed and evaluated that might integrate these findings with a mathematical model of the interpersonal influence system that alters some or all of its members' positions on an issue. In this paper we provide evidence that a general model in the network science on opinion dynamics substantially clarifies how truth wins in groups.

network science | opinion dynamics | truth statements

Max Planck's famous null hypothesis on truth propagation is that "a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die." Classic experiments in social psychology have probed the conditions of "truth wins" in social groups. This evidence presents mixed findings, and no dynamical system model has been advanced that integrates them. Thorndike and Barnlund (1, 2) show that an initial consensus is adopted without scrutiny whether or not it is true. Asch (3–5) shows that individuals are likely to abandon the truth if they are located in a group in which all other members have a fixed agreement on a false position. Sherif (6) shows that a group converges to a false consensus when all its members are uncertain about the true position. Laughlin and Adamopoulos (7) and Laughlin and Ellis (8) find that reaching consensus on a true position depends on the size of the faction of individuals with the correct position on an issue. Clearly, truth propagation is not a viral contagion in which a true position is automatically adopted by anyone who comes into contact with it, and the hazard rate of adopting false positions on issues is substantial.

The above referenced findings of the classical literature on truth wins in social groups are based on statistical tests of hypotheses on input–output associations and do not specify the groups' throughput mechanisms that alter individuals' positions on issues. Understanding the mechanisms by which truth wins or fails to win is an open problem. It is a problem that has become increasingly important with the development of social media that allow a rapid widespread dissemination of true and false information. Unfortunately, we are currently confronted with such a hazard in which a faction of policy makers have concluded that the numerical evidence on global warming does not indicate any trajectory toward a dangerous threshold level that would require immediate action.

Investigations of the conditions of truth wins have not included a network science modeling of the opinion dynamics that determine groups' outcomes. We present evidence that an attention to such dynamics substantially clarifies how truth wins in groups. Given information on a group's initial opinions and influence network on a quantitative issue, network science models of opinion dynamics are motivated by the idea that complex sequences of opinion updates are triggered by individuals' expo-

sure to a set of one or more other individuals' fixed or changing positions on an issue. An interdisciplinary field of research has developed on the problem of predicting group members' displayed settled positions on an issue. Such predictions require a mathematical model of interpersonal influence systems in which repetitive individual-level updates of quantitative opinions are occurring. Numerous models of such systems have been proposed; see refs. 9 and 10 for reviews. Consistent with all mathematically based science on dynamical natural systems, the goal of this field is to find an empirically supported parsimonious general mechanism of opinion updating that unfolds on any influence network topology and, thus, explains a large domain of observed realizations of influence system outcomes. The open fundamental question is whether or not there exists an elementary "cognitive algebra" (11, 12) in individuals' integration of their own and others' displayed opinions. The most widely accepted candidate is a weighted averaging mechanism with which each individual automatically allocates weights to his or her own and others' fixed or changing opinions to obtain an updated opinion that is a convex combination of opinions. The set of all group members' allocated weights defines the influence network of the group. Thus, a group issue-specific influence network is automatically assembled as an implication of individuals' information integration mechanism; that is, the weights are in the mechanism used by individuals to adjust their own opinion on an issue. In this literature, the Friedkin–Johnsen (13, 14) generalization of the seminal French–Harary–DeGroot (15–17) weighted-averaging opinion update mechanism is currently the only model on which a sustained line of human-subject experiments (14, 18–20) have been conducted to evaluate predictions of opinion changes. This model is given by

$$x_i(k+1) = a_{ii} \sum_{j=1}^n w_{ij} x_j(k) + (1 - a_{ii}) x_i(0), \forall i, \quad [1]$$

Significance

Investigations of the conditions of truth wins in social groups present mixed findings and have not included a network science modeling of the opinion dynamics that alter individuals' positions. This paper presents evidence that an attention to such dynamics substantially clarifies how truth wins. The findings of our experiments on groups of human subjects suggest that the mixed findings on truth wins in social groups can be integrated as special-case realizations of a general model of opinion dynamics on influence networks.

Author contributions: N.E.F. and F.B. designed research; N.E.F. performed research; N.E.F. analyzed data; and N.E.F. and F.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence should be addressed. Email: friedkin@soc.ucsb.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1710603114/-DCSupplemental.

$k = 0, 1, \dots$, where $x_i(0) \in \mathcal{R}$, $0 \leq w_{ij} \leq 1 \forall ij$, $\sum_{j=1}^n w_{ij} = 1$, and $a_{ii} = 1 - w_{ii} \forall i$. The corresponding matrix equation for the influence system is

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{W}\mathbf{x}(k) + (\mathbf{I} - \mathbf{A})\mathbf{x}(0), \quad k = 0, 1, \dots, \quad [2]$$

where \mathbf{W} is the system's $n \times n$ matrix of w_{ij} weights and \mathbf{A} is the system's $n \times n$ diagonal matrix of individuals' levels of stubborn attachment to their initial opinions ($0 \leq a_{ii} \leq 1 \forall i$, $a_{ij} = 0 \forall i \neq j$). In the special case of $\mathbf{A} = \mathbf{I}$, the model reduces to a French–Harary–DeGroot system without resistances to opinion change. Mathematical analysis has shown that this system's equilibrium equation,

$$\mathbf{x}(\infty) = \mathbf{A}\mathbf{W}\mathbf{x}(\infty) + (\mathbf{I} - \mathbf{A})\mathbf{x}(0), \quad [3]$$

is consistent with relaxations of the model's simplifying assumptions of synchronous opinion updates (21, 22) and time-invariant \mathbf{W} (23) that allow random asynchronous updates and time-varying weights. The model provides theoretical foundations for widely used measures of the influence centralities of individuals in a group: In the special case of an $\mathbf{A} = \mathbf{I}$, its measure is equivalent to eigenvector centrality, and in the special case of an $\mathbf{A} = \alpha \mathbf{I}$, $0 < \alpha < 1$, its measure is equivalent to PageRank centrality (24–26). See [SI Text](#) on other relevant properties of this model. Experiments on groups of human subjects have evaluated the model's predictive accuracy on quantitative issues of judgment for which there are no true or false numerical positions. The designs of these experiments include groups in face-to-face interaction (20) and groups in which interaction is constrained by different telephonic communication structures that prohibit direct conversations among particular pairs of individuals and allow random dyadic conversations of varying length (14, 18, 19). Its predictive accuracy includes exact predictions of observed final opinions such as in the following case that appears in the findings of the experiments reported in this paper. Before the group discussion, the four members of this group privately recorded their initial opinions,

$$\mathbf{x}(0) = \begin{bmatrix} 65 & 74 & 74 & 52 \end{bmatrix}^T.$$

A direct measure of the cognitive w_{ij} weights of Eq. 1 (relating to the way an individual is integrating information in his or her own mind) is based on the four individuals' independent postdiscussion reports of the extent to which each of the other group members influenced their own final opinions; see [SI Text](#) on this measure. In this group on this issue, 1 reports according no weight to itself or to 4, and equal weights to 2 and 3, and the group's matrix of such reports is:

$$\mathbf{W} = \begin{bmatrix} 0 & 0.50 & 0.50 & 0 \\ 0 & 0.25 & 0.50 & 0.25 \\ 0.15 & 0.30 & 0.30 & 0.25 \\ 0 & 0.50 & 0.50 & 0 \end{bmatrix}.$$

The model's predicted equilibrium opinions are

$$\begin{aligned} \mathbf{x}(\infty) &= \mathbf{A}\mathbf{W}\mathbf{x}(\infty) + (\mathbf{I} - \mathbf{A})\mathbf{x}(0), \\ &= (\mathbf{I} - \mathbf{A}\mathbf{W})^{-1}(\mathbf{I} - \mathbf{A})\mathbf{x}(0). \end{aligned}$$

Let $\mathbf{V} = (\mathbf{I} - \mathbf{A}\mathbf{W})^{-1}(\mathbf{I} - \mathbf{A})$. Its v_{ij} elements correspond to the total relative (direct and indirect) influence of group member j 's initial opinion on the final opinion of group member i . Each $0 \leq v_{ij} \leq 1 \forall ij$ and $\sum_{j=1}^n v_{ij} = 1 \forall i$. Thus, for this group, the model's predicted final opinions are

$$\hat{\mathbf{x}}(\infty) = \mathbf{V}\mathbf{x}(0) = \begin{bmatrix} 74 & 74 & 74 & 74 \end{bmatrix}^T,$$

and, in this case, they are exact predictions of the individuals' observed privately reported final opinions. We will show that the model's predictive accuracy is high in the pool of all

cases reported in this article on intellectual (truth statement) issues. Note that there is nothing in this prediction that indicates whether any opinion is correct or incorrect.

The fact that the issue is intellectual, in contrast to the judgment issues that have been involved in all previously conducted experiments, is irrelevant to the model's explanation of the observed opinion changes. Here, and in general, the foundations of this model's success rests (i) on its assumption of a weighted averaging mechanism of opinion updates that allows for levels of stubborn attachment to initial opinions and (ii) on an equilibrium equation that is robust to violations of its simplifying assumptions of synchronous updates on a time-invariant influence network.

However, there are important implications of quantitative intellectual issues that distinguish them from quantitative judgment issues. Note that in the above example two individuals reported identical initial opinions. Henceforth, we refer to a subset of individuals with identical initial opinions as a faction. Such factions frequently occur on the intellectual issues of the conducted experiments. In contrast, in previous experiments on quantitative judgment issues (14, 19, 20), initial opinion factions rarely occur so that typically any initial opinion in a group is a distinct position that is associated with one individual. On the issues investigated in this paper, because the same initial position is often held by two or more individuals the displayed position of a faction-as-a-unit has an influence value. This idea has motivated the classical work on truth wins and abandonment referenced in the Introduction. We break from the tradition of treating the relative sizes of factions as an explanatory variable. Consistent with a network science approach, we define the influence of a faction-as-a-unit on any individual i as the sum of the faction members' v_{ij} total influences on i . Thus, one member of the faction may explain most its influence, and other individuals may be more or less influenced by different members of a faction. On this basis, we can formulate a prediction of the relative total influence of the truth, that is a value in the $[0, 1]$ interval, in determining each i 's final opinion on an issue as follows:

$$c_i = \sum_{j=1}^n v_{ij} u_j(0), \quad u_j(0) = \begin{cases} 1, & \text{if } x_j(0) \text{ is true} \\ 0, & \text{if } x_j(0) \text{ is false} \end{cases} \quad [4]$$

For each individual i , regardless of whether his or her own initial opinion is true or false, the hypothesis is that $0 \leq c_i \leq 1$ predicts whether i will hold a true or false final opinion. For each group, the hypothesis is that $0 \leq \frac{1}{n} \sum_{i=1}^n c_i \leq 1$ predicts whether a group consensus is reached on a true or false position. This formulation allows for cases in which a true initial position is abandoned, or a true initial opinion of one member alters all other members' false opinions, or the true initial position of a faction alters or fails to alter other members' opinions. We will show that our findings support this network science approach for explaining how truth wins in groups.

In addition, we show that a deeper level of persuasion is involved in group discussions of truth statements. In the classic Asch (3–5) and Sherif (6) experiments, individuals voiced their positions without discussion. Group discussion allows not only a display of opinions but also a display of the calculative logics upon which different initial opinions are based. The abandonment of an opinion may be associated with the adoption of a true or false calculative logic that applies to the issue, and to all other issues to which the logic can be applied. This deeper level of persuasion, based on social learning, allows a growth of initial opinion factions along a sequence of issues on which a common calculative logic can be adapted to each issue. In the conducted experiments, the structure of each issue varies over the issue sequence so that formulaic rote learning may fail in the absence of a more abstract understanding of the logic involved. Although individuals may have been persuaded to accept a true opinion, they may present either a true or false initial opinion

on the next issue of the issue sequence, depending on whether they acquired an abstract understanding of how to obtain a true opinion. Similarly, they may be persuaded to accept a false opinion without acquiring an abstract understanding of its calculative logic. We show that both true and false initial opinion factions grow in size along a sequence of intellective issues on which a true or false calculative logic can be adapted to each issue. In our data, faction growth includes cases of group trajectories along an issue sequence to extremal states of initial consensus. In these extremal cases, all group members have the same initial opinion (true or false) and maximal levels of attachment to it. Finally, we consider the net resultant of these intimately intertwined processes of interpersonal influence and social learning. When there are various false calculative logics and only one true calculative logic, the distribution of initial opinion errors (for all individuals nested in independent groups) evolves over time along an issue sequence toward a distribution in which the frequency of true initial opinions dominates. Although the bulk of initial opinions may be based on false calculative logics, they are group-specific. Thus, truth wins in a decentralized population of individuals nested in independent groups under the condition of one-true versus many-false calculative logics on a class of intellective issues.

We conducted experiments on risk evaluation issues—analytical reliability problems—for which there are objectively correct numerical positions. We assembled 45 face-to-face groups of three to four individuals. Each group dealt with a random permutation of the same issues under the instruction that they should try to reach a consensus on each issue. With 161 individuals nested in 45 groups with three to four members, each dealing with a sequence of eight issues, the experiments set up 360 group-issue occasions in which individuals present an initial and final opinion and 1,288 individual-level behavioral observations of (initial, final) tuples of opinion. The individuals were recruited from the undergraduate subject pool of our university and paid for their participation. The University of California, Santa Barbara Institutional Review Board approved the study, and all subjects provided written informed consent. The experiment is a standard prepost group-discussion randomized design. One of the eight issues is posed to a group. (i) Individuals privately record their initial positions on the issue, (ii) a group discussion is then opened and concluded, and (iii) individuals pri-

vately record their final positions and their subjective weights indicating the extent to which each of the other group members influenced their own final opinion. Then another issue is posed, and so on, until the concluding issue of the random permutation of the eight issues is reached. For each group-specific issue, the obtained data are the group's n -vectors of initial and final opinions and its $n \times n$ matrix of allocated weights, which are measures of the model's $\mathbf{x}(0)$, $\mathbf{x}(\infty)$, and \mathbf{W} constructs, respectively. See *SI Text* on the detailing of these experiments. Because the data structure is a longitudinal multilevel design in which individuals are nested in groups, the statistical analysis draws on methods suitable to such clustered data (*SI Text*). We also conducted a baseline experiment on 30 individuals who worked independently on random permutations of the same eight issues.

Results

First, we evaluate the predictions of the opinion dynamics model. The experiments provided 1,288 individual-level behavioral observations of (initial, final) tuples of opinion. Fig. 1 evaluates the model's predicted final opinions. The Pearson product-moment correlation coefficient (ρ) of observed and predicted final opinions is $\rho = 0.939$ ($P < 0.001$, $n = 1,288$).

Opinion changes were prevalent, and the model's predictions of the direction and magnitude of these changes are associated with the observed changes $\rho = 0.873$ ($P < 0.001$). Regressing individuals' observed final opinions on their predicted final opinions and observed initial opinions, the estimated coefficient for the predicted values is 0.983 ($P < 0.001$, $SE = 0.041$) and the coefficient for initial opinions is insignificant ($P = 0.496$). These findings on quantitative intellective issues are strongly consistent with previous findings on quantitative judgment issues (14, 19, 20). The network science speculation that there may be an elementary mechanism of opinion updates, based on weighted averaging, is buttressed by these findings. With respect to the specialized literature of experiments on groups dealing with intellective problems, referenced in the Introduction of this paper, these findings suggest that a network science model of opinion dynamics may serve to integrate this literature's observations. The special cases of this model cover all realizations of initial opinions (initial consensus, faction structures, and complete heterogeneity), levels of stubborn attachments to initial opinions,

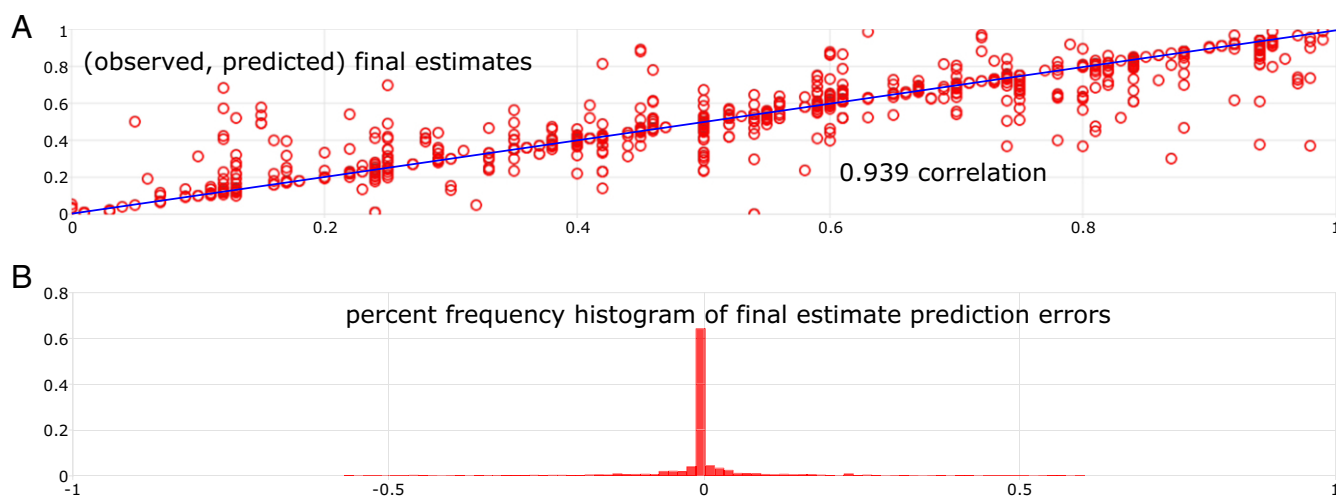


Fig. 1. The F-J model of group influence systems is applied to 161 individuals nested in 45 groups, each dealing with a sequence of eight issues, setting up 360 group-task occasions in which the individuals nested in a group presented an initial and final opinion and 1,288 individual-level behavioral observations. (A) Individuals' observed and F-J model-predicted final opinions. The Pearson product-moment correlation of the observed and predicted final opinions is 0.939, $P < 0.001$. (B) Percent frequency histogram of the F-J model prediction errors of final estimates.

and influence network topologies. In the present experiments we have not constrained these realizations as in the classic Asch (3–5) experiments and have allowed them to naturally arise.

Second, we evaluate the Eq. 4 hypothesis that the probability of an individual's adopting a true final opinion on an issue depends on whether one or more true initial opinions exist in a group and their influence on an individual's opinion. We find that the hazard of a false final opinion given a false initial opinion is (i) near 100% if an individual is not exposed to a true initial opinion and (ii) reduced but not eliminated (in these data to 18%) if an individual is exposed to a true initial opinion. The existence of a true initial opinion in a group is necessary but not sufficient to explain its adoption (who adopts it or who does not). The Eq. 4 hypothesis is that its adoption by a particular individual depends on the existence of one or more true initial opinions and on the influence of the one or more individuals who are advocating it. In our data, a group's array of initial opinions is not constrained by the experimental design as in the classic Asch (3–5) experiments on truth abandonment, where a single naive subject with a true initial opinion is confronted with two or more confederates of the experimenter who display an identical false fixed opinion on each issue of a sequence of similar issues. We allow true or false factions of initial opinions to naturally occur along an issue sequence, and we place no design constraints on individuals' levels of stubbornness. In the 360 occasions that present arrays of initial opinions of true or false factions, all feasible faction structures of initial opinion were observed. In the 152 group-issue-specific occasions involving triads, the distribution of the number of unique initial opinions is 40, 50, and 62 occasions with one, two, and three unique initial opinions, respectively. In the 208 group-issue-specific occasions involving tetrads, the distribution of the number of unique initial opinions is 34, 63, 54, and 57 occasions with one, two, three, and four unique initial opinions, respectively. The existence of initial opinion factions presents an open question in the network science on opinion dynamics. Can the effects of initial opinion factions on group members' final opinions be understood as a simple aggregation of the relative total interpersonal influences of the individuals who are advocating a particular position on an issue? We address this question with a logistic regression of the $n = 1,288$ individuals' observed binary adoptions of true versus false final opinions on the Eq. 4 predicted influence value of the truth on their final opinions, controlling for the truth state of their observed initial opinions. The estimated coefficient for the predicted influence of the truth is 7.747 ($P < 0.001$, $SE = 1.040$) and the estimated coefficient of

the truth state of their initial opinions is insignificant ($P = 0.552$). Thus, Eq. 4 captures instances of maintained or adjusted initial-to-final opinions and predicts the log odds of adopting true final positions. Since $0 \leq c_i \leq 1$ and its observed values span its feasible values, the estimated 7.747 coefficient indicates a strong association (linear increases in the log odds of adopting a true final opinion up to a nearly eight-fold increase) consistent with the predictions of the entertained opinion dynamics model.

For each group, the hypothesis is that the group's mean $0 \leq \frac{1}{n} \sum_{i=1}^n c_i \leq 1$ predicts whether a group consensus is reached on a true position. From the opinion dynamics model, this mean value must be 0 if all group members have false initial opinions or if any true initial opinion has no influence on final opinions, and its positive values indicate the extent to which one or more true initial opinions in a group influence final opinions. The data present 360 occasions on which a consensus might be reached. We evaluate the hypothesis with a logistic regression of the observed events of a true versus false final consensus on the mean Eq. 4 predicted total influence of the truth on final opinions, controlling for the number of group members with true initial opinions. The estimated coefficient for the predicted influence of the truth is 6.888 ($P < 0.001$, $SE = 1.651$) and the estimated coefficient for the number of group members with true initial opinions is insignificant ($P = 0.317$). Since $0 \leq \frac{1}{n} \sum_{i=1}^n c_i \leq 1$ and its observed values span its feasible values, the estimated 6.888 coefficient indicates a strong association (linear increases in the log odds of a true final consensus up to a nearly sevenfold increase) consistent with the predictions of the entertained opinion dynamics model. The insignificant coefficient for the number of group members with true initial opinions is consistent with the hypothesis that the opinion dynamics model explains the association of initial factions of true positions on true consensus outcomes.

Third, we show that a deeper level of persuasion is involved in group discussions of truth statements. Group discussion allows not only a display of opinions but also a display of the calculative logics upon which different initial opinions are based. The abandonment of an opinion may be associated with adoption of a true or false calculative logic that applies to the issue, and to all other issues to which the logic can be applied. This deeper level of persuasion, based on social learning, allows a growth of initial opinion factions along a sequence of issues on which a common calculative logic can be adapted to each issue. On the analytical reliability issues of this experiment, the structure of each issue varies over the issue sequence so that formulaic rote learning

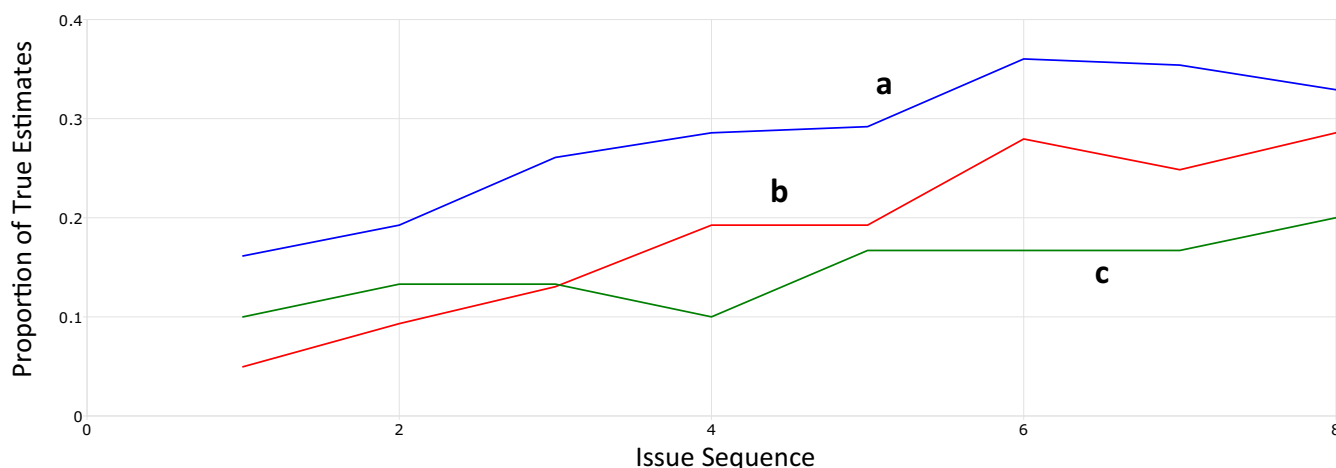


Fig. 2. The proportion true opinions along the issue sequence for (a) the final opinions of the 161 individuals nested in groups, (b) the initial opinions of the 161 individuals nested in groups, and (c) the opinions of the 30 isolated individuals.

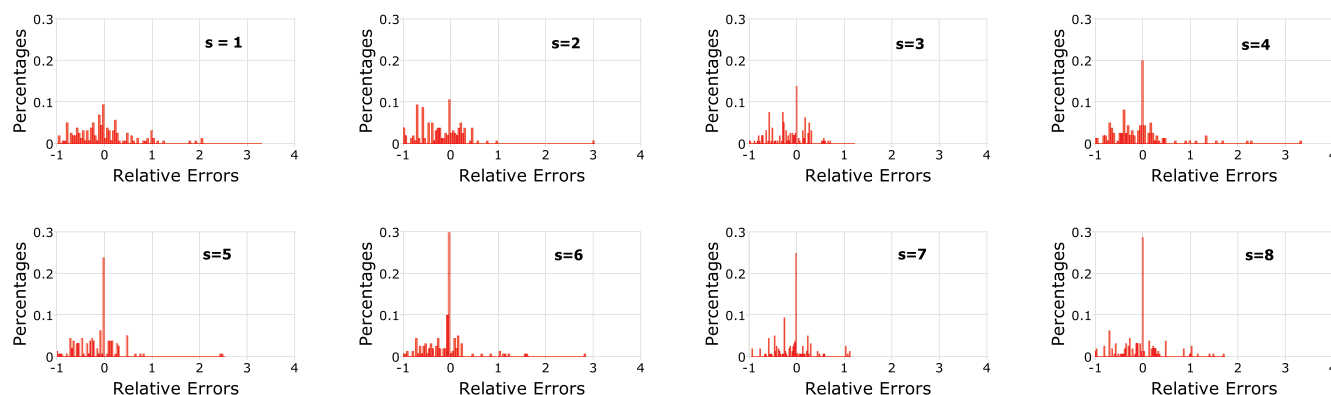


Fig. 3. The emergence of a dominant faction of individuals with true initial opinions in the sample of 161 individuals nested in 45 disconnected groups.

may fail in the absence of a more abstract understanding of the logic involved. Although individuals may have been persuaded to accept a true opinion, they may present either a true or false initial opinion on the next issue of the issue sequence, depending on whether they acquired an abstract understanding of how to obtain a true opinion. Similarly, they may be persuaded to accept a false opinion without acquiring an abstract understanding of its calculative logic. Thus, true and false social learning is manifested in the growth of initial opinion factions along an issue sequence.

Fig. 2 reports the proportion of true opinions along the issue sequence for 30 isolated individuals and the 161 individuals nested in 45 groups. The effect of groups is the difference between the (a) curve for the final opinions of individuals nested in groups and the (c) curve for the isolated individuals. The (b) curve for the initial opinions of the individuals nested in groups signals the occurrence of social learning events that alter the conditions under which the opinion dynamics of the groups unfold. The difference of the (a) and (b) curves for final and initial opinions is based on the additional factor of interpersonal systems in which the influence of one or more true initial opinions elevates the adoption of a true final opinion whether or not the adopter understands how such opinions were obtained. Note that the maximum proportion of true final opinions on the (a) curve is below 0.40. A substantial proportion of the groups in these data fail to present any true opinion. In such groups, factions of initial opinion grow along the issue sequence based on false learning.

Finally, we consider the net resultant of these intimately intertwined processes of interpersonal influence and social learning. When there are various false calculative logics and only one true calculative logic, the resulting distribution of individuals' initial opinion errors is an evolving distribution in which the frequency of true initial opinions is elevated. Fig 3 shows the evolving distribution of relative initial opinion errors. Truth wins in a decentralized population of individuals nested in independent groups that are dealing with a sequence of intellctive issues on which a common calculative logic can be adapted to each issue. Such evolution will occur when the application of incorrect facts and logic is group-idiosyncratic and, thus, fails to form any substantial between-group faction with a common error bias. Such evolution may be derailed by social movements or social media that elevate the adoption of a particular set of false facts and logic.

Discussion

The research reported in this paper suggests that a network science of opinion dynamics may advance our understanding of how truth wins in groups. It presents an approach that contrasts with the literature on the wisdom of crowds (27–30) in which true

numerical estimates are obtained from the mean value of the distribution of independent individual estimates. This work also contrasts with the literature on groupthink (31–35) with its implication that the definitional properties of groupthink are not necessarily associated with flawed decisions. Both positive and negative groupthink outcomes are permissible. We have noted that a group's array of initial opinions evolves along an issue sequence and includes cases of group trajectories to extremal states of both true or false initial consensus based on the adoption of a particular calculative logic. If the shared calculative logic is false, then an instance of negative groupthink is expressed. If the shared calculative logic is true, then an instance of positive groupthink is expressed. Formally, both outcomes are consistent with the conditions of groupthink. Moreover, this work contributes a way of thinking about how truth wins in a population of independent groups. Social science has long emphasized the importance shared norms in large-scale populations (6, 36, 37), because shared norms appear to be the only basis on which independent groups might have similar initial opinions on the same class of issues, and reach a similar consensus on each issue. It is for this reason that shared norms have been emphasized by social scientists as a basis of societal coherence. We have shown how a dominant faction with opinions based on a true calculative logic, for a class of issues, may emerge in a population of independent groups. Although the bulk of initial opinions in such a population may be based on false calculative logics, they are group-specific. Truth wins in a decentralized population of individuals nested in independent groups under the condition of one-true versus many-false calculative logics. Its dominance is based on the property that any individual who understands the relevant science or mathematics must come to the same conclusion, while those who do not understand the relevant science or mathematics are likely to adopt diverse group idiosyncratic conclusions. This one-true versus many-false calculative logics condition of truth wins is not satisfied when social movements or social media elevate the adoption of a particular set of false facts and logic. Macro-level belief system dynamics that generate shared constraints on individuals' opinions are complex (26, 38) and not easily controlled. Both history and social science present the evidence that the hazard rate of adopting false positions on truth statements is substantial and that truth propagation is not a viral contagion in which a true position is automatically adopted by anyone who comes into contact with it. Incentivizing correct opinions, to lower this hazard rate, has not been investigated. Fact checks that are exogenous feedbacks to groups would incentivize a rethinking of the method being used to obtain problem solutions. However, the conditions affecting the willingness or capacity of groups to alter their agreed-upon methods of solving problems are currently not well understood. More

broadly, the advancement of a network science modeling of the group dynamics that alter individuals' positions on issues is currently limited by the absence of a technology that provides more direct evidence on the causal basis of these dynamics and, in particular, a more direct measurement of the weights in the automatic convex combination mechanism that is the

widely assumed cognitive algebra of iterative opinion update processes.

ACKNOWLEDGMENTS. We thank Wenjun Mei, Victor Amelkin, and Sose Abraamyan for their helpful contributions to this work and Janelle Pham, who administered the data collection. This work was supported by US Army Research Laboratory and US Army Research Office Grant W911NF-15-1-0577.

1. Thorndike RL (1938) The effect of discussion upon the correctness of decisions, when the factor of majority influence is allowed for. *J Soc Psychol* 9:343–363.
2. Barnlund CD (1959) A comparative study of individual, majority, and group judgment. *J Abnorm Psychol* 58:55–60.
3. Asch SE (1951) Effects of group pressure upon the modification and distortion of judgments. *Groups, Leadership, and Men*, ed Guetzkow H (Carnegie, Pittsburgh), pp 177–190.
4. Asch S (1952) *Social Psychology* (Prentice Hall, Englewood Cliffs, NJ).
5. Asch SE (1956) Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychol Monogr Gen Appl* 70:1–70.
6. Sherif M (1936) *The Psychology of Norms* (Harper, New York).
7. Laughlin PR, Adamopoulos J (1980) Social combination processes and individual learning for six-person cooperative groups on an intellectual task. *J Pers Soc Psychol* 38:941–947.
8. Laughlin PR, Ellis AL (1986) Demonstrability and social combination processes on mathematical intellectual tasks. *J Exp Soc Psychol* 22:177–189.
9. Proskurnikov AV, Tempo R (2017) A tutorial on modeling and analysis of dynamic social networks. Part I. *Annu Rev Contr* 43:65–79.
10. Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. *Rev Mod Phys* 81:591–646.
11. Anderson NH (1971) Integration theory and attitude change. *Psychol Rev* 78:171–206.
12. Anderson NH (1982) Cognitive algebra and social psychophysics. *Social Attitudes and Psychophysical Measurement*, ed Wegener B (Lawrence Erlbaum Associates, Hillsdale, NJ), pp 123–148.
13. Friedkin NE, Johnsen EC (1990) Social influence and opinions. *J Math Sociol* 15:193–206.
14. Friedkin NE, Johnsen EC (1999) Social influence networks and opinion change. *Adv Group Process* 16:1–29.
15. French JRP (1956) A formal theory of social power. *Psychol Rev* 63:181–194.
16. Harary F (1959) A criterion for unanimity in French's theory of social power. *Studies in Social Power*, ed Cartwright D (Institute for Social Research, Univ of Michigan, Ann Arbor, MI), pp 168–182.
17. DeGroot MH (1974) Reaching a consensus. *J Am Stat Assoc* 69:118–121.
18. Friedkin NE, Cook KS (1990) Peer group influence. *Socio Meth Res* 19:122–143.
19. Friedkin NE, Johnsen EC (2011) *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics* (Cambridge Univ Press, New York).
20. Friedkin NE, Jia P, Bullo F (2016) A theory of the evolution of social power: Natural trajectories of interpersonal influence systems along issue sequences. *Sociol Sci* 3:444–472.
21. Ravazzi C, Frasca P, Tempo R, Ishii H (2015) Ergodic randomized algorithms and dynamics over networks. *IEEE Trans Control Netw Syst* 2:78–87.
22. Frasca P, Ravazzi C, Tempo R, Ishii H (2013) Gossips and prejudices: Ergodic randomized dynamics in social networks. *IFAC Proc Vol* 46:212–219.
23. Proskurnikov AV, Tempo R, Cao M, Friedkin NE (2017) Opinion evolution in time-varying social influence networks with prejudiced agents. arXiv:1704.06900.
24. Friedkin NE (1991) Theoretical foundations for centrality measures. *Am J Sociol* 96:1478–1504.
25. Friedkin NE, Johnsen EC (2014) Two steps to obfuscation. *Soc Network* 39:12–13.
26. Parsegov SE, Proskurnikov AV, Tempo R, Friedkin NE (2016) Novel multidimensional models of opinion dynamics in social networks. *IEEE Trans Automat Contr* 62:2270–2285.
27. Galton F (1907) Vox populi (the wisdom of crowds). *Nature* 75:450–451.
28. Surowiecki J (2004) *The Wisdom of Crowds* (Anchor, New York).
29. Golub B, Jackson MO (2010) Naive learning in social networks and the wisdom of crowds. *Am Econ J Microecon* 2:112–149.
30. Lorenz J, Rauhut H, Schweitzer F, Helbing D (2011) How social influence can undermine the wisdom of crowd effect. *Proc Natl Acad Sci USA* 108:9020–9025.
31. Janis IL (1972) *Victims of Groupthink* (Houghton Mifflin, Boston).
32. Janis IL (1989) *Crucial Decisions* (Free, New York).
33. McCauley C (1989) The nature of social influence in groupthink: Compliance and internalization. *J Pers Soc Psychol* 57:250–240.
34. McCauley C (1998) Group dynamics in Janis's theory of groupthink: Backward and forward. *Organ Behav Hum Decis Process* 73:142–162.
35. Park W-W (2000) A comprehensive empirical investigation of the relationships among variables of the groupthink model. *J Organ Behav* 21:873–887.
36. Parsons T (1951) *The Social System* (Routledge, New York).
37. Hechter M, Opp K-D (2001) *Social Norms* (Russell Sage Foundation, New York).
38. Friedkin NE, Proskurnikov AV, Tempo R, Parsegov SE (2016) Network science on belief system dynamics under logic constraints. *Science* 354:321–326.