

## RESEARCH

## SOCIAL SCIENCE

# The spread of true and false news online

Soroush Vosoughi,<sup>1</sup> Deb Roy,<sup>1</sup> Sinan Aral<sup>2\*</sup>

We investigated the differential diffusion of all of the verified true and false news stories distributed on Twitter from 2006 to 2017. The data comprise ~126,000 stories tweeted by ~3 million people more than 4.5 million times. We classified news as true or false using information from six independent fact-checking organizations that exhibited 95 to 98% agreement on the classifications. Falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information, and the effects were more pronounced for false political news than for false news about terrorism, natural disasters, science, urban legends, or financial information. We found that false news was more novel than true news, which suggests that people were more likely to share novel information. Whereas false stories inspired fear, disgust, and surprise in replies, true stories inspired anticipation, sadness, joy, and trust. Contrary to conventional wisdom, robots accelerated the spread of true and false news at the same rate, implying that false news spreads more than the truth because humans, not robots, are more likely to spread it.

**F**oundational theories of decision-making (1–3), cooperation (4), communication (5), and markets (6) all view some conceptualization of truth or accuracy as central to the functioning of nearly every human endeavor. Yet, both true and false information spreads rapidly through online media. Defining what is true and false has become a common political strategy, replacing debates based on a mutually agreed on set of facts. Our economies are not immune to the spread of falsity either. False rumors have affected stock prices and the motivation for large-scale investments, for example, wiping out \$130 billion in stock value after a false tweet claimed that Barack Obama was injured in an explosion (7). Indeed, our responses to everything from natural disasters (8, 9) to terrorist attacks (10) have been disrupted by the spread of false news online.

New social technologies, which facilitate rapid information sharing and large-scale information cascades, can enable the spread of misinformation (i.e., information that is inaccurate or misleading). But although more and more of our access to information and news is guided by these new technologies (11), we know little about their contribution to the spread of falsity online. Though considerable attention has been paid to anecdotal analyses of the spread of false news by the media (12), there are few large-scale empirical investigations of the diffusion of misinformation or its social origins. Studies of the spread of misinformation are currently limited to analyses of small, ad hoc samples that ignore two of the most important scientific questions: How do truth and falsity diffuse differently, and what factors of human judgment explain these differences?

Current work analyzes the spread of single rumors, like the discovery of the Higgs boson (13) or the Haitian earthquake of 2010 (14), and multiple rumors from a single disaster event, like the Boston Marathon bombing of 2013 (10), or it develops theoretical models of rumor diffusion (15), methods for rumor detection (16), credibility evaluation (17, 18), or interventions to curtail the spread of rumors (19). But almost no studies comprehensively evaluate differences in the spread of truth and falsity across topics or examine why false news may spread differently than the truth. For example, although Del Vicario *et al.* (20) and Bessi *et al.* (21) studied the spread of scientific and conspiracy-theory stories, they did not evaluate their veracity. Scientific and conspiracy-theory stories can both be either true or false, and they differ on stylistic dimensions that are important to their spread but orthogonal to their veracity. To understand the spread of false news, it is necessary to examine diffusion after differentiating true and false scientific stories and true and false conspiracy-theory stories and controlling for the topical and stylistic differences between the categories themselves. The only study to date that segments rumors by veracity is that of Fricker *et al.* (19), who analyzed ~4000 rumors spreading on Facebook and focused more on how fact checking affects rumor propagation than on how falsity diffuses differently than the truth (22).

In our current political climate and in the academic literature, a fluid terminology has arisen around “fake news,” foreign interventions in U.S. politics through social media, and our understanding of what constitutes news, fake news, false news, rumors, rumor cascades, and other related terms. Although, at one time, it may have been appropriate to think of fake news as referring to the veracity of a news story, we now believe that this phrase has been irredeemably polarized in our current political and media climate. As politicians have implemented a political strategy of labeling news sources that do not

support their positions as unreliable or fake news, whereas sources that support their positions are labeled reliable or not fake, the term has lost all connection to the actual veracity of the information presented, rendering it meaningless for use in academic classification. We have therefore explicitly avoided the term fake news throughout this paper and instead use the more objectively verifiable terms “true” or “false” news. Although the terms fake news and misinformation also imply a willful distortion of the truth, we do not make any claims about the intent of the purveyors of the information in our analyses. We instead focus our attention on veracity and stories that have been verified as true or false.

We also purposefully adopt a broad definition of the term news. Rather than defining what constitutes news on the basis of the institutional source of the assertions in a story, we refer to any asserted claim made on Twitter as news (we defend this decision in the supplementary materials section on “reliable sources,” section S1.2). We define news as any story or claim with an assertion in it and a rumor as the social phenomena of a news story or claim spreading or diffusing through the Twitter network. That is, rumors are inherently social and involve the sharing of claims between people. News, on the other hand, is an assertion with claims, whether it is shared or not.

A rumor cascade begins on Twitter when a user makes an assertion about a topic in a tweet, which could include written text, photos, or links to articles online. Others then propagate the rumor by retweeting it. A rumor’s diffusion process can be characterized as having one or more cascades, which we define as instances of a rumor-spreading pattern that exhibit an unbroken retweet chain with a common, singular origin. For example, an individual could start a rumor cascade by tweeting a story or claim with an assertion in it, and another individual could independently start a second cascade of the same rumor (pertaining to the same story or claim) that is completely independent of the first cascade, except that it pertains to the same story or claim. If they remain independent, they represent two cascades of the same rumor. Cascades can be as small as size one (meaning no one retweeted the original tweet). The number of cascades that make up a rumor is equal to the number of times the story or claim was independently tweeted by a user (not retweeted). So, if a rumor “A” is tweeted by 10 people separately, but not retweeted, it would have 10 cascades, each of size one. Conversely, if a second rumor “B” is independently tweeted by two people and each of those two tweets is retweeted 100 times, the rumor would consist of two cascades, each of size 100.

Here we investigate the differential diffusion of true, false, and mixed (partially true, partially false) news stories using a comprehensive data set of all of the fact-checked rumor cascades that spread on Twitter from its inception in 2006 to 2017. The data include ~126,000 rumor cascades spread by ~3 million people more than 4.5 million times. We sampled all rumor cascades investigated by six independent fact-checking organizations

<sup>1</sup>Massachusetts Institute of Technology (MIT), the Media Lab, E14-526, 75 Amherst Street, Cambridge, MA 02142, USA. <sup>2</sup>MIT, E62-364, 100 Main Street, Cambridge, MA 02142, USA.

\*Corresponding author. Email: sinan@mit.edu

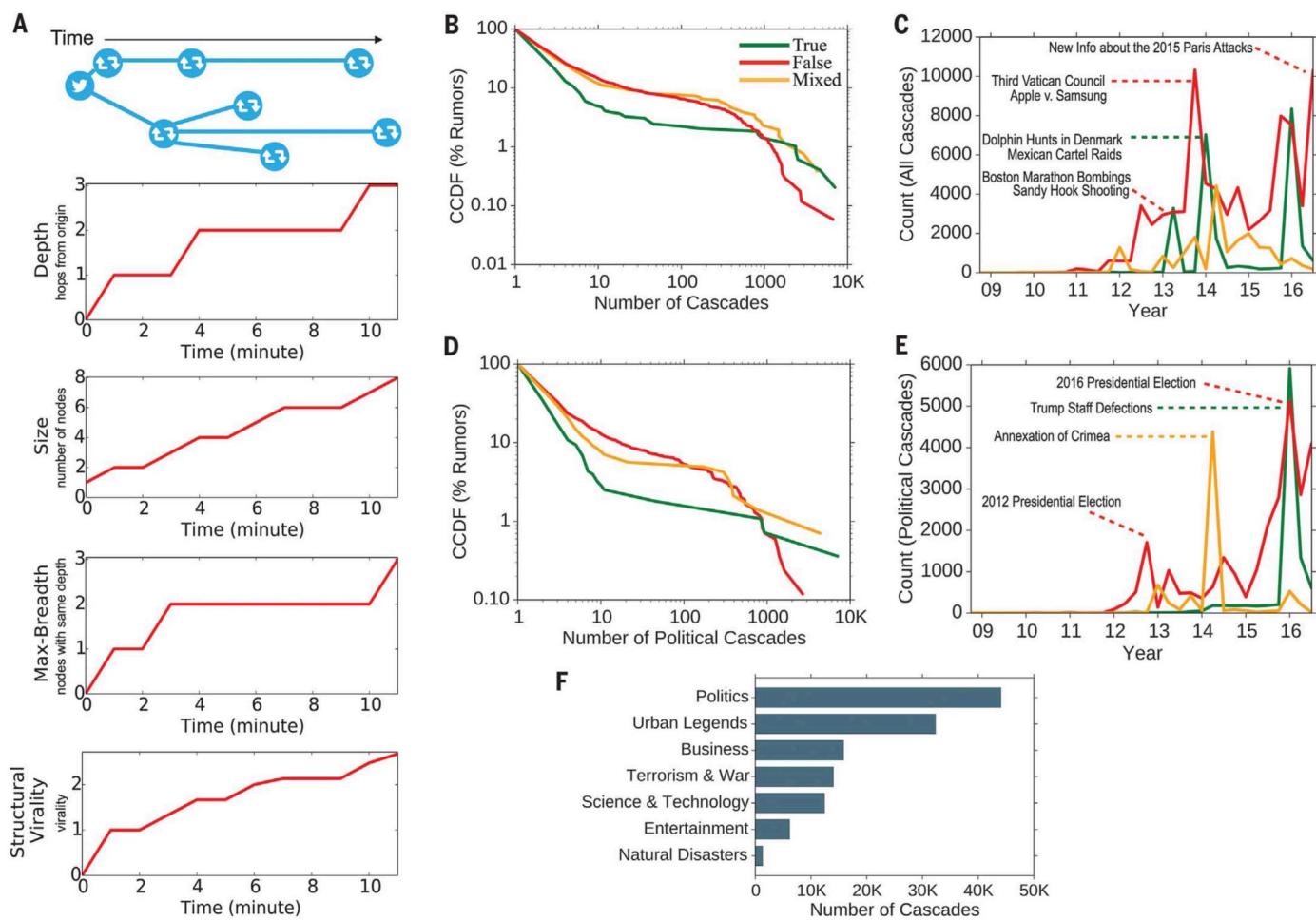
(snopes.com, politifact.com, factcheck.org, truthfiction.com, hoax-slayer.com, and urbanlegends.about.com) by parsing the title, body, and verdict (true, false, or mixed) of each rumor investigation reported on their websites and automatically collecting the cascades corresponding to those rumors on Twitter. The result was a sample of rumor cascades whose veracity had been agreed on by these organizations between 95 and 98% of the time. We cataloged the diffusion of the rumor cascades by collecting all English-language replies to tweets that contained a link to any of the aforementioned websites from 2006 to 2017 and used optical character recognition to extract text from images where needed. For each reply tweet, we extracted the original tweet being replied to and all the retweets of the original tweet. Each retweet cascade represents a rumor propagating on Twitter that has been verified as true or false by the fact-checking organizations (see the supplementary materials for more details on cascade construction). We then quantified the cascades'

depth (the number of retweet hops from the origin tweet over time, where a hop is a retweet by a new unique user), size (the number of users involved in the cascade over time), maximum breadth (the maximum number of users involved in the cascade at any depth), and structural virality (23) (a measure that interpolates between content spread through a single, large broadcast and that which spreads through multiple generations, with any one individual directly responsible for only a fraction of the total spread) (see the supplementary materials for more detail on the measurement of rumor diffusion).

As a rumor is retweeted, the depth, size, maximum breadth, and structural virality of the cascade increase (Fig. 1A). A greater fraction of false rumors experienced between 1 and 1000 cascades, whereas a greater fraction of true rumors experienced more than 1000 cascades (Fig. 1B); this was also true for rumors based on political news (Fig. 1D). The total number of false rumors peaked at the end of both 2013 and 2015 and again at the

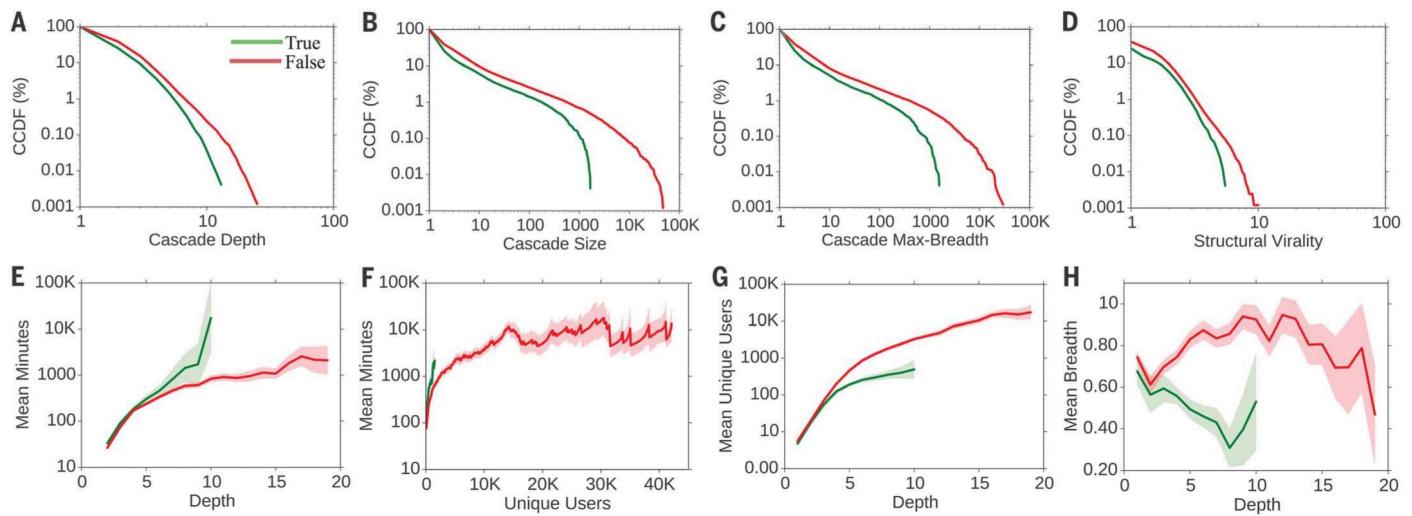
end of 2016, corresponding to the last U.S. presidential election (Fig. 1C). The data also show clear increases in the total number of false political rumors during the 2012 and 2016 U.S. presidential elections (Fig. 1E) and a spike in rumors that contained partially true and partially false information during the Russian annexation of Crimea in 2014 (Fig. 1E). Politics was the largest rumor category in our data, with ~45,000 cascades, followed by urban legends, business, terrorism, science, entertainment, and natural disasters (Fig. 1F).

When we analyzed the diffusion dynamics of true and false rumors, we found that falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information [Kolmogorov-Smirnov (K-S) tests are reported in tables S3 to S10]. A significantly greater fraction of false cascades than true cascades exceeded a depth of 10, and the top 0.01% of false cascades diffused eight hops deeper into the Twittersphere than the truth, diffusing to depths



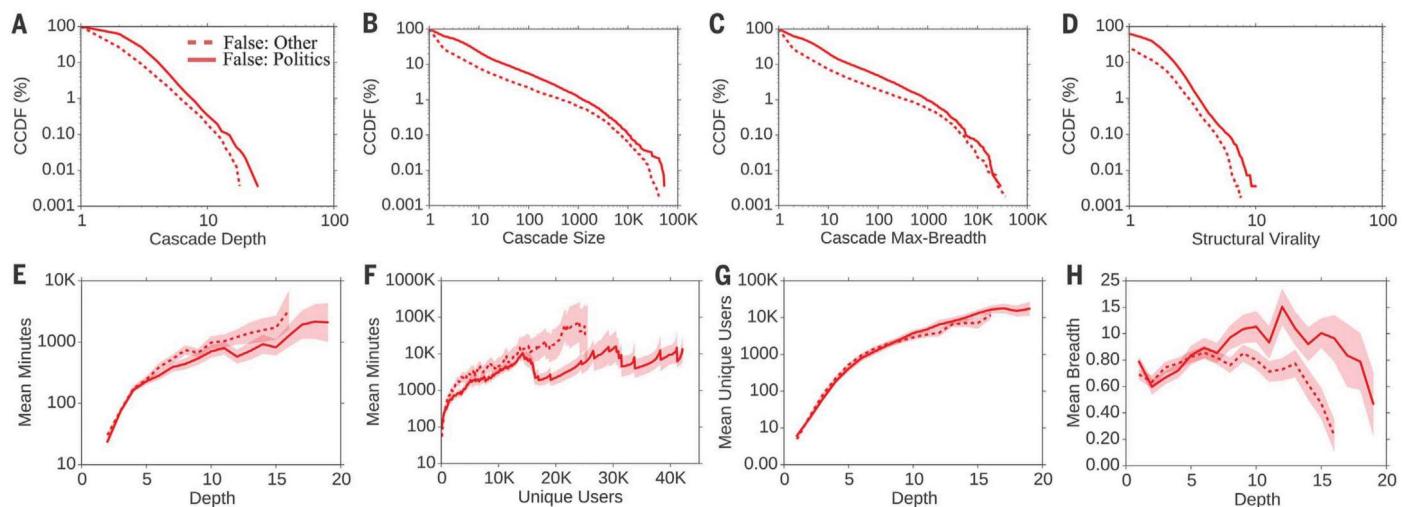
**Fig. 1. Rumor cascades.** (A) An example rumor cascade collected by our method as well as its depth, size, maximum breadth, and structural virality over time. “Nodes” are users. (B) The complementary cumulative distribution functions (CCDFs) of true, false, and mixed (partially true and partially false) cascades, measuring the fraction of rumors that exhibit a given number of cascades. (C) Quarterly counts of all true, false, and mixed rumor cascades

that diffused on Twitter between 2006 and 2017, annotated with example rumors in each category. (D) The CCDFs of true, false, and mixed political cascades. (E) Quarterly counts of all true, false, and mixed political rumor cascades that diffused on Twitter between 2006 and 2017, annotated with example rumors in each category. (F) A histogram of the total number of rumor cascades in our data across the seven most frequent topical categories.



**Fig. 2. Complementary cumulative distribution functions (CCDFs) of true and false rumor cascades.** (A) Depth. (B) Size. (C) Maximum breadth. (D) Structural virality. (E and F) The number of minutes it takes for true and false rumor cascades to reach any (E) depth and (F) number of unique Twitter users. (G) The number of unique Twitter

users reached at every depth and (H) the mean breadth of true and false rumor cascades at every depth. In (H), plot is lognormal. Standard errors were clustered at the rumor level (i.e., cascades belonging to the same rumor were clustered together; see supplementary materials for additional details).



**Fig. 3. Complementary cumulative distribution functions (CCDFs) of false political and other types of rumor cascades.** (A) Depth. (B) Size. (C) Maximum breadth. (D) Structural virality. (E and F) The number of minutes it takes for false political and other false news cascades to reach

any (E) depth and (F) number of unique Twitter users. (G) The number of unique Twitter users reached at every depth and (H) the mean breadth of these false rumor cascades at every depth. In (H), plot is lognormal. Standard errors were clustered at the rumor level.

greater than 19 hops from the origin tweet (Fig. 2A). Falsehood also reached far more people than the truth. Whereas the truth rarely diffused to more than 1000 people, the top 1% of false-news cascades routinely diffused to between 1000 and 100,000 people (Fig. 2B). Falsehood reached more people at every depth of a cascade than the truth, meaning that many more people retweeted falsehood than they did the truth (Fig. 2C). The spread of falsehood was aided by its virality, meaning that falsehood did not simply spread through broadcast dynamics but rather through peer-to-peer diffusion characterized by a viral branching process (Fig. 2D).

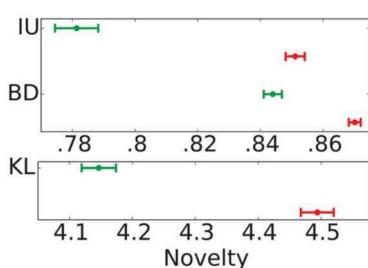
It took the truth about six times as long as falsehood to reach 1500 people (Fig. 2F) and 20 times as long as falsehood to reach a cascade depth of 10 (Fig. 2E). As the truth never diffused beyond a depth of 10, we saw that falsehood reached a depth of 19 nearly 10 times faster than the truth reached a depth of 10 (Fig. 2E). Falsehood also diffused significantly more broadly (Fig. 2H) and was retweeted by more unique users than the truth at every cascade depth (Fig. 2G).

False political news (Fig. 1D) traveled deeper (Fig. 3A) and more broadly (Fig. 3C), reached more people (Fig. 3B), and was more viral than any other category of false information (Fig. 3D). False po-

litical news also diffused deeper more quickly (Fig. 3E) and reached more than 20,000 people nearly three times faster than all other types of false news reached 10,000 people (Fig. 3F). Although the other categories of false news reached about the same number of unique users at depths between 1 and 10, false political news routinely reached the most unique users at depths greater than 10 (Fig. 3G). Although all other categories of false news traveled slightly more broadly at shallower depths, false political news traveled more broadly at greater depths, indicating that more-popular false political news items exhibited broader and more-accelerated diffusion dynamics

**A**

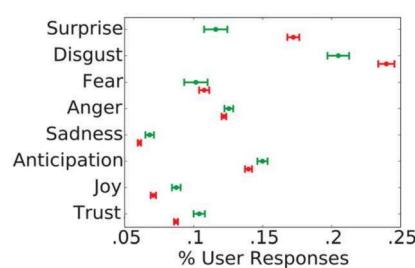
	median false—true	mean false—true	mean (log) false—true	stdv (log) false—true	ks-test
followers	410	466	2234	5240	D=0.104, p~0.0
followees	383	509	1002	1707	D=0.136, p~0.0
verified	0	0	0.002	0.006	D=0.005, p<0.001
engagement	9.52	9.54	19.70	24.65	D=0.054, p~0.0
account age	982	1214	1072	1269	D=0.125, p~0.0

**C****E**

	mean false—true	variance false—true	ks-test
IU	0.85	0.78	D=0.457, p~0.0
KL	4.49	4.15	D=0.433, p~0.0
BD	0.87	0.84	D=0.415, p~0.0

**B**

	coeff	odds ratio	std err	z	P> z	[95.0% Conf. Int.]
account age	0.0002	1.000160	2.07e-05	7.759	0.000	0.000 0.000
engagement	0.0066	1.006648	0.000	18.019	0.000	0.006 0.007
falsehood	0.5350	1.707489	0.084	6.366	0.000	0.370 0.700
followees	-1.639e-05	0.999984	8.73e-06	-1.877	0.060	-3.35e-05 7.22e-07
followers	5.192e-05	1.000052	7.77e-06	6.682	0.000	3.67e-05 6.72e-05
intercept	-2.3941	0.091257	0.072	-33.437	0.000	-2.534 -2.254
verified	1.4261	4.162467	0.090	15.915	0.000	1.250 1.602

**D****F**

	mean false—true	variance false—true	ks-test
surprise	0.172	0.116	D=0.205, p~0.0
disgust	0.240	0.205	D=0.102, p~0.0
fear	0.108	0.102	D=0.021, p~0.164
anger	0.122	0.126	D=0.023, p~0.078
sadness	0.061	0.068	D=0.037, p~0.0
anticipation	0.140	0.150	D=0.038, p~0.0
joy	0.071	0.087	D=0.061, p~0.0
trust	0.087	0.104	D=0.060, p~0.0

**Fig. 4. Models estimating correlates of news diffusion, the novelty of true and false news, and the emotional content of replies to news.**

(A) Descriptive statistics on users who participated in true and false rumor cascades as well as K-S tests of the differences in the distributions of these measures across true and false rumor cascades. (B) Results of a logistic regression model estimating users' likelihood of retweeting a rumor as a function of variables shown at the left. coeff, logit coefficient; z, z score. (C) Differences in the information uniqueness (IU), scaled Bhattacharyya distance (BD), and K-L divergence (KL) of true (green) and false (red) rumor tweets compared to the corpus of prior tweets the user was exposed to in the 60 days before retweeting the rumor tweet. (D) The emotional

(Fig. 3H). Analysis of all news categories showed that news about politics, urban legends, and science spread to the most people, whereas news about politics and urban legends spread the fastest and were the most viral in terms of their structural virality (see fig. S11 for detailed comparisons across all topics).

One might suspect that structural elements of the network or individual characteristics of the users involved in the cascades explain why falsity travels with greater velocity than the truth. Perhaps those who spread falsity “followed” more people, had more followers, tweeted more often, were more often “verified” users, or had been on Twitter longer. But when we compared users involved in true and false rumor cascades, we found that the opposite was true in every case. Users who spread false news had significantly fewer followers (K-S test = 0.104,  $P \sim 0.0$ ), followed significantly fewer people (K-S test = 0.136,  $P \sim 0.0$ ), were significantly less active on Twitter (K-S test = 0.054,  $P \sim 0.0$ ), were verified significantly less often (K-S test = 0.004,  $P < 0.001$ ), and had been on Twitter for significantly less time (K-S test = 0.125,  $P \sim 0.0$ ) (Fig. 4A). Falsehood

diffused farther and faster than the truth despite these differences, not because of them.

When we estimated a model of the likelihood of retweeting, we found that falsehoods were 70% more likely to be retweeted than the truth (Wald chi-square test,  $P \sim 0.0$ ), even when controlling for the account age, activity level, and number of followers and followees of the original tweeter, as well as whether the original tweeter was a verified user (Fig. 4B). Because user characteristics and network structure could not explain the differential diffusion of truth and falsity, we sought alternative explanations for the differences in their diffusion dynamics.

One alternative explanation emerges from information theory and Bayesian decision theory. Novelty attracts human attention (24), contributes to productive decision-making (25), and encourages information sharing (26) because novelty updates our understanding of the world. When information is novel, it is not only surprising, but also more valuable, both from an information theoretic perspective [in that it provides the greatest aid to decision-making (25)] and from a social perspective [in that it conveys so-

content of replies to true (green) and false (red) rumor tweets across seven dimensions categorized by the NRC. (E) Mean and variance of the IU, KL, and BD of true and false rumor tweets compared to the corpus of prior tweets the user has seen in the 60 days before seeing the rumor tweet as well as K-S tests of their differences across true and false rumors. (F) Mean and variance of the emotional content of replies to true and false rumor tweets across seven dimensions categorized by the NRC as well as K-S tests of their differences across true and false rumors. All standard errors are clustered at the rumor level, and all models are estimated with cluster-robust standard errors at the rumor level.

cial status on one that is “in the know” or has access to unique “inside” information (26)]. We therefore tested whether falsity was more novel than the truth and whether Twitter users were more likely to retweet information that was more novel.

To assess novelty, we randomly selected ~5000 users who propagated true and false rumors and extracted a random sample of ~25,000 tweets that they were exposed to in the 60 days prior to their decision to retweet a rumor. We then specified a latent Dirichlet Allocation Topic model (27), with 200 topics and trained on 10 million English-language tweets, to calculate the information distance between the rumor tweets and all the prior tweets that users were exposed to before retweeting the rumor tweets. This generated a probability distribution over the 200 topics for each tweet in our data set. We then measured how novel the information in the true and false rumors was by comparing the topic distributions of the rumor tweets with the topic distributions of the tweets to which users were exposed in the 60 days before their retweet. We found that false rumors were significantly more

novel than the truth across all novelty metrics, displaying significantly higher information uniqueness ( $K\text{-}S$  test = 0.457,  $P \sim 0.0$ ) (28), Kullback-Leibler ( $K\text{-}L$ ) divergence ( $K\text{-}S$  test = 0.433,  $P \sim 0.0$ ) (29), and Bhattacharyya distance ( $K\text{-}S$  test = 0.415,  $P \sim 0.0$ ) (which is similar to the Hellinger distance) (30). The last two metrics measure differences between probability distributions representing the topical content of the incoming tweet and the corpus of previous tweets to which users were exposed.

Although false rumors were measurably more novel than true rumors, users may not have perceived them as such. We therefore assessed users' perceptions of the information contained in true and false rumors by comparing the emotional content of replies to true and false rumors. We categorized the emotion in the replies by using the leading lexicon curated by the National Research Council Canada (NRC), which provides a comprehensive list of ~140,000 English words and their associations with eight emotions based on Plutchik's (31) work on basic emotion—anger, fear, anticipation, trust, surprise, sadness, joy, and disgust (32)—and a list of ~32,000 Twitter hashtags and their weighted associations with the same emotions (33). We removed stop words and URLs from the reply tweets and calculated the fraction of words in the tweets that related to each of the eight emotions, creating a vector of emotion weights for each reply that summed to one across the emotions. We found that false rumors inspired replies expressing greater surprise ( $K\text{-}S$  test = 0.205,  $P \sim 0.0$ ), corroborating the novelty hypothesis, and greater disgust ( $K\text{-}S$  test = 0.102,  $P \sim 0.0$ ), whereas the truth inspired replies that expressed greater sadness ( $K\text{-}S$  test = 0.037,  $P \sim 0.0$ ), anticipation ( $K\text{-}S$  test = 0.038,  $P \sim 0.0$ ), joy ( $K\text{-}S$  test = 0.061,  $P \sim 0.0$ ), and trust ( $K\text{-}S$  test = 0.060,  $P \sim 0.0$ ) (Fig. 4, D and F). The emotions expressed in reply to falsehoods may illuminate additional factors, beyond novelty, that inspire people to share false news. Although we cannot claim that novelty causes retweets or that novelty is the only reason why false news is retweeted more often, we do find that false news is more novel and that novel information is more likely to be retweeted.

Numerous diagnostic statistics and manipulation checks validated our results and confirmed their robustness. First, as there were multiple cascades for every true and false rumor, the variance of and error terms associated with cascades corresponding to the same rumor will be correlated. We therefore specified cluster-robust standard errors and calculated all variance statistics clustered at the rumor level. We tested the robustness of our findings to this specification by comparing analyses with and without clustered errors and found that, although clustering reduced the precision of our estimates as expected, the directions, magnitudes, and significance of our results did not change, and chi-square ( $P \sim 0.0$ ) and deviance ( $d$ ) goodness-of-fit tests ( $d = 3.4649 \times 10^{-6}$ ,  $P \sim 1.0$ ) indicate that the models are well specified (see supplementary materials for more detail).

Second, a selection bias may arise from the restriction of our sample to tweets fact checked by the six organizations we relied on. Fact checking may select certain types of rumors or draw additional attention to them. To validate the robustness of our analysis to this selection and the generalizability of our results to all true and false rumor cascades, we independently verified a second sample of rumor cascades that were not verified by any fact-checking organization. These rumors were fact checked by three undergraduate students at Massachusetts Institute of Technology (MIT) and Wellesley College. We trained the students to detect and investigate rumors with our automated rumor-detection algorithm running on 3 million English-language tweets from 2016 (34). The undergraduate annotators investigated the veracity of the detected rumors using simple search queries on the web. We asked them to label the rumors as true, false, or mixed on the basis of their research and to discard all rumors previously investigated by one of the fact-checking organizations. The annotators, who worked independently and were not aware of one another, agreed on the veracity of 90% of the 13,240 rumor cascades that they investigated and achieved a Fleiss' kappa of 0.88. When we compared the diffusion dynamics of the true and false rumors that the annotators agreed on, we found results nearly identical to those estimated with our main data set (see fig. S17). False rumors in the robustness data set had greater depth ( $K\text{-}S$  test = 0.139,  $P \sim 0.0$ ), size ( $K\text{-}S$  test = 0.131,  $P \sim 0.0$ ), maximum breadth ( $K\text{-}S$  test = 0.139,  $P \sim 0.0$ ), structural virality ( $K\text{-}S$  test = 0.066,  $P \sim 0.0$ ), and speed (fig. S17) and a greater number of unique users at each depth (fig. S17). When we broadened the analysis to include majority-rule labeling, rather than unanimity, we again found the same results (see supplementary materials for results using majority-rule labeling).

Third, although the differential diffusion of truth and falsity is interesting with or without robot, or bot, activity, one may worry that our conclusions about human judgment may be biased by the presence of bots in our analysis. We therefore used a sophisticated bot-detection algorithm (35) to identify and remove all bots before running the analysis. When we added bot traffic back into the analysis, we found that none of our main conclusions changed—false news still spread farther, faster, deeper, and more broadly than the truth in all categories of information. The results remained the same when we removed all tweet cascades started by bots, including human retweets of original bot tweets (see supplementary materials, section S8.3) and when we used a second, independent bot-detection algorithm (see supplementary materials, section S8.3.5) and varied the algorithm's sensitivity threshold to verify the robustness of our analysis (see supplementary materials, section S8.3.4). Although the inclusion of bots, as measured by the two state-of-the-art bot-detection algorithms we used in our analysis, accelerated the spread of both true and false news, it affected their spread roughly equally. This suggests that false

news spreads farther, faster, deeper, and more broadly than the truth because humans, not robots, are more likely to spread it.

Finally, more research on the behavioral explanations of differences in the diffusion of true and false news is clearly warranted. In particular, more robust identification of the factors of human judgment that drive the spread of true and false news online requires more direct interaction with users through interviews, surveys, lab experiments, and even neuroimaging. We encourage these and other approaches to the investigation of the factors of human judgment that drive the spread of true and false news in future work.

False news can drive the misallocation of resources during terror attacks and natural disasters, the misalignment of business investments, and misinformed elections. Unfortunately, although the amount of false news online is clearly increasing (Fig. 1, C and E), the scientific understanding of how and why false news spreads is currently based on ad hoc rather than large-scale systematic analyses. Our analysis of all the verified true and false rumors that spread on Twitter confirms that false news spreads more pervasively than the truth online. It also overturns conventional wisdom about how false news spreads. Though one might expect network structure and individual characteristics of spreaders to favor and promote false news, the opposite is true. The greater likelihood of people to retweet falsity more than the truth is what drives the spread of false news, despite network and individual factors that favor the truth. Furthermore, although recent testimony before congressional committees on misinformation in the United States has focused on the role of bots in spreading false news (36), we conclude that human behavior contributes more to the differential spread of falsity and truth than automated robots do. This implies that misinformation-containment policies should also emphasize behavioral interventions, like labeling and incentives to dissuade the spread of misinformation, rather than focusing exclusively on curtailing bots. Understanding how false news spreads is the first step toward containing it. We hope our work inspires more large-scale research into the causes and consequences of the spread of false news as well as its potential cures.

## REFERENCES AND NOTES

1. L. J. Savage, *J. Am. Stat. Assoc.* **46**, 55–67 (1951).
2. H. A. Simon, *The New Science of Management Decision* (Harper & Brothers Publishers, New York, 1960).
3. R. Wedgwood, *Noûs* **36**, 267–297 (2002).
4. E. Fehr, U. Fischbacher, *Nature* **425**, 785–791 (2003).
5. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379–423 (1948).
6. S. Bikhchandani, D. Hirshleifer, I. Welch, *J. Polit. Econ.* **100**, 992–1026 (1992).
7. K. Rapoza, "Can 'fake news' impact the stock market?" *Forbes*, 26 February 2017; www.forbes.com/sites/kenrapoza/2017/02/26/can-fake-news-impact-the-stock-market/.
8. M. Mendoza, B. Poblete, C. Castillo, in *Proceedings of the First Workshop on Social Media Analytics* (Association for Computing Machinery, ACM, 2010), pp. 71–79.
9. A. Gupta, H. Lamba, P. Kumaraguru, A. Joshi, in *Proceedings of the 22nd International Conference on World Wide Web* (ACM, 2010), pp. 729–736.

10. K. Starbird, J. Maddock, M. Orand, P. Achterman, R. M. Mason, in *iConference 2014 Proceedings* (iSchools, 2014).
11. J. Gottfried, E. Shearer, "News use across social media platforms," Pew Research Center, 26 May 2016; [www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/](http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/).
12. C. Silverman, "This analysis shows how viral fake election news stories outperformed real news on Facebook," *BuzzFeed News*, 16 November 2016; [www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook/](http://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook/).
13. M. De Domenico, A. Lima, P. Mougel, M. Musolesi, *Sci. Rep.* **3**, 2980 (2013).
14. O. Oh, K. H. Kwon, H. R. Rao, in *Proceedings of the International Conference on Information Systems* (International Conference on Information Systems, ICIS, paper 231, 2010).
15. M. Tamburino, G. Ruffo, A. Flammini, F. Menczer, in *Proceedings of the 24th International Conference on World Wide Web* (ACM, 2015), pp. 977–982.
16. Z. Zhao, P. Resnick, Q. Mei, in *Proceedings of the 24th International Conference on World Wide Web* (ACM, 2015), pp. 1395–1405.
17. M. Gupta, P. Zhao, J. Han, in *Proceedings of the 2012 Society for Industrial and Applied Mathematics International Conference on Data Mining* (Society for Industrial and Applied Mathematics, SIAM, 2012), pp. 153–164.
18. G. L. Ciampaglia et al., *PLOS ONE* **10**, e0128193 (2015).
19. A. Friggeri, L. A. Adamic, D. Eckles, J. Cheng, in *Proceedings of the International Conference on Weblogs and Social Media* (Association for the Advancement of Artificial Intelligence, AAAI, 2014).
20. M. Del Vicario et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 554–559 (2016).
21. A. Bessi et al., *PLOS ONE* **10**, e0118093 (2015).
22. Friggeri et al. (19) do evaluate two metrics of diffusion: depth, which shows little difference between true and false rumors, and shares per rumor, which is higher for true rumors than it is for false rumors. Although these results are important, they are not definitive owing to the smaller sample size of the study; the early timing of the sample, which misses the rise of false news after 2013; and the fact that more shares per rumor do not necessarily equate to deeper, broader, or more rapid diffusion.
23. S. Goel, A. Anderson, J. Hofman, D. J. Watts, *Manage. Sci.* **62**, 180–196 (2015).
24. L. Itti, P. Baldi, *Vision Res.* **49**, 1295–1306 (2009).
25. S. Aral, M. Van Alstyne, *Am. J. Sociol.* **117**, 90–171 (2011).
26. J. Berger, K. L. Milkman, *J. Mark. Res.* **49**, 192–205 (2012).
27. D. M. Blei, A. Y. Ng, M. I. Jordan, *J. Mach. Learn. Res.* **3**, 993–1022 (2003).
28. S. Aral, P. Dhillon, "Unpacking novelty: The anatomy of vision advantages," Working paper, MIT-Sloan School of Management, Cambridge, MA, 22 June 2016; [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2388254](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2388254).
29. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley, 2012).
30. T. Kailath, *IEEE Trans. Commun. Technol.* **15**, 52–60 (1967).
31. R. Plutchik, *Am. Sci.* **89**, 344–350 (2001).
32. S. M. Mohammad, P. D. Turney, *Comput. Intell.* **29**, 436–465 (2013).
33. S. M. Mohammad, S. Kiritchenko, *Comput. Intell.* **31**, 301–326 (2015).
34. S. Vosoughi, D. Roy, in *Proceedings of the 10th International AAAI Conference on Weblogs and Social Media* (AAAI, 2016), pp. 707–710.
35. C. A. Davis, O. Varol, E. Ferrara, A. Flammini, F. Menczer, in *Proceedings of the 25th International Conference Companion on World Wide Web* (ACM, 2016), pp. 273–274.
36. For example, this is an argument made in recent testimony by Clint Watts—Robert A. Fox Fellow at the Foreign Policy

Research Institute and Senior Fellow at the Center for Cyber and Homeland Security at George Washington University—given during the U.S. Senate Select Committee on Intelligence hearing on "Disinformation: A Primer in Russian Active Measures and Influence Campaigns" on 30 March 2017; [www.intelligence.senate.gov/sites/default/files/documents/os-cwatts-033017.pdf](http://www.intelligence.senate.gov/sites/default/files/documents/os-cwatts-033017.pdf).

## ACKNOWLEDGMENTS

We are indebted to Twitter for providing funding and access to the data. We are also grateful to members of the MIT research community for invaluable discussions. The research was approved by the MIT institutional review board. The analysis code is freely available at <https://goo.gl/forms/AKIIZujpxehN7Y33>. The entire data set is also available, from the same link, upon signing an access agreement stating that (i) you shall only use the data set for the purpose of validating the results of the MIT study and for no other purpose; (ii) you shall not attempt to identify, reidentify, or otherwise deanonymize the data set; and (iii) you shall not further share, distribute, publish, or otherwise disseminate the data set. Those who wish to use the data for any other purposes can contact and make a separate agreement with Twitter.

## SUPPLEMENTARY MATERIALS

[www.science.org/content/359/6380/1146/suppl/DC1](http://www.science.org/content/359/6380/1146/suppl/DC1)  
Materials and Methods  
Figs. S1 to S20  
Tables S1 to S39  
References (37–75)

14 September 2017; accepted 19 January 2018  
10.1126/science.aap9559