



PROJECT MUSE®

---

## Fake News and Partisan Epistemology

Regina Rini

Kennedy Institute of Ethics Journal, Volume 27, Number 2 Supplement, June 2017, pp. E-43-E-64 (Article)

Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/ken.2017.0025>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/670860>

## Fake News and Partisan Epistemology

**ABSTRACT.** This paper does four things: (1) It provides an analysis of the concept ‘fake news.’ (2) It identifies distinctive epistemic features of social media testimony. (3) It argues that partisanship-in-testimony-reception is not always epistemically vicious; in fact some forms of partisanship are consistent with individual epistemic virtue. (4) It argues that a solution to the problem of fake news will require changes to institutions, such as social media platforms, not just to individual epistemic practices.

**D**id you know that Hillary Clinton sold weapons to ISIS? Or that Mike Pence called Michelle Obama “the most vulgar First Lady we’ve ever had”? No, you didn’t know these things. You couldn’t know them, because these claims are false.<sup>1</sup> But many American voters believed them.

One of the most distinctive features of the 2016 campaign was the rise of “fake news,” factually false claims circulated on social media, usually via channels of partisan camaraderie. Media analysts and social scientists are still debating what role fake news played in Trump’s victory.<sup>2</sup> But whether or not it drove the outcome, fake news certainly affected the choices of some individual voters.

Why were people willing to believe easily dis-confirmable, often ridiculous, stories? In this paper I will suggest the following answer: people believe fake news because they acquire it through social media sharing, which is a peculiar sort of testimony. Social media sharing has features that reduce audience willingness to think critically or check facts. This effect is amplified when the testifier and audience share a partisan orientation. Shared partisan affiliation encourages testimony recipients to grant more credibility to testifiers than would otherwise be warranted.

So far these points may seem familiar. But the deeper aim of this paper is to normatively evaluate how fake news is transmitted, and here my

answer may be less expected. I will argue that fake news transmission is often *individually reasonable*. That is, individual people typically act reasonably when they grant greater credibility to fellow partisans, even if this sometimes leads to the acquisition of false beliefs. This normative analysis generates a further claim about the remedy for fake news: it will not be solved by focusing on individual epistemic virtue. Rather, we must treat fake news as a tragedy of the epistemic commons, and its solution as a coordination problem. Fake news exploits otherwise reasonable practices of information transmission. Ending it will require institutional change.

This paper has four goals, corresponding to the following four parts. First, I give an analysis of the concept ‘fake news.’ Second, I identify the unusual epistemic features of news transmission via social media testimony. Third, I argue that partisanship in testimony is sometimes individually reasonable, and can be consistent with epistemic virtue despite predictably generating false beliefs. Fourth, I argue that we should treat the harms of partisan epistemology as an institutional, rather than individual, problem, and I offer an example of institutional improvement.

#### WHAT IS FAKE NEWS?

What is fake news? It is not merely false information conveyed by reportage. As the word ‘fake’ suggests, fake news requires *intentional deception*; honest reporting errors are not fake news.<sup>3</sup> When *TIME* journalist Zeke Miller falsely reported on Twitter that Donald Trump had removed a bust of MLK Jr. from the Oval Office, this was not fake news. Miller mistakenly believed that the bust *had* been removed, though in fact it was merely hidden behind a door (Gibbs 2017). Fake news is not merely false; it is deceptive.

But this is not yet sufficient to fully characterize fake news, because fake news involves a *particular type* of deception. It is more than mere lying. Suppose you ask me why I did not come to your Jeff Sessions Appreciation Party, and I falsely claim that I was doing the laundry. In fact, I was undergoing a superfluous root canal, an experience I deemed preferable to attending your party. This is not fake news, though obviously it is intentional deception. The ‘news’ part of ‘fake news’ implies that the deception is intended for an audience larger than the immediate recipient; fake news is meant to be shared and shared again.

The intentions behind fake news are also more complicated than in simple cases of lying. A moment ago I said that fake news requires intentional deception, but this may be too strong. Deception is not always

the primary goal of fake news. Often the motive is financial rather than epistemic. Entire businesses, now infamously concentrated in Macedonia, exist to generate fake news headlines that attract fervent Internet clicking—and ad revenue (Silverman and Alexander 2016). The film studio 20th Century Fox recently created websites full of click-bait fake stories in order to attract social media attention to promote a new film (Rainey 2017). Presumably these entrepreneurial Macedonian teenagers and film producers did not care whether anyone ended up believing their fake news, so long as the clicks kept coming.

These examples show the complexity of motives for fake news; spreading false information is not the only goal. But deception does play *some* role, even in these cases. Fake news works as click-bait only if a large number of people choose to share links, and presumably this requires that at least some of them believe the story. People who make money from fake news are perfectly happy if nine-in-ten of their readers are not deceived, but they *do* need some percentage to be deceived long enough to convey the link to future clickers. So we can say that creators of fake news intend to deceive at least *a part* of their overall audience, even if this deception is merely instrumental and not the ultimate goal.

Of course, other fake news creators *do* intend to deceive as many people as possible. Committed partisans try to erode their opponents' support by tricking persuadable voters. Foreign actors may be involved as well; some analysts claim that anti-Clinton fake news was manufactured by shady groups with links to Russian military intelligence.<sup>4</sup> For these creators, fake news needs to travel widely not only to generate clicks, but also to change epistemic states. We can call this aimed-at-deception form 'pure' fake news, while also keeping in mind the impure, deception-as-instrument form motivated by financial gain.

So, we can finally give a clear definition of fake news. A fake news story is one that purports to describe events in the real world, typically by mimicking the conventions of traditional media reportage, yet is known by its creators to be significantly false, and is transmitted with the two goals of being widely re-transmitted and of deceiving at least some of its audience.<sup>5</sup>

You'll note that my definition of fake news does not specify *how* it is transmitted. In particular, I have not specified that it is spread through social media. Fake news can be spread other ways—email chains, posters on streetlamps, etc. But there is a strong contingent relationship between fake news and social media, especially in the 2016 election. I will therefore focus on social media fake news.

## THE BENT TESTIMONY OF SOCIAL MEDIA

Why do people believe fake news? A first-pass answer is easy enough; they believe fake news because it is presented to them via testimony, and like most of us they typically accept testimony from others, all else equal. Fake news stories turn up in their social media feeds, evidently endorsed by people whom they trust (to some degree), and it's natural to believe what trusted friends tell you.

The epistemology of testimony has received significant attention from philosophers in recent decades (e.g., Coady 1992; Lackey 2008; Goldberg 2010). A person counts as believing a proposition on the basis of testimony when she believes it *because* the proposition was presented to her by another person. This is typically an epistemically virtuous practice, as we rely upon others for our knowledge of many things distant from us in space or time. A community of people with a practice of accepting one another's testimony will be able to learn far more than individuals who insist upon believing only what they discover on their own.

Of course, an epistemic practice of uncritically accepting testimony would be prime for abuse by liars and bullshitters.<sup>6</sup> Sensible use of testimony requires norms for blocking the acceptance of suspect cases. Some of these norms have to do with the identity of the testifier or features of her current motivation.<sup>7</sup> It is wise to suspend default acceptance of testimony from someone who wishes to sell you a used car. Other norms are about testimonial content; like all sources of evidence, it is reasonable to suspend confidence in a piece of testimony if it is radically at odds with what you already know about how the world works. If a new acquaintance tells me that she saw a squirrel steal a park-goer's slice of pizza, I'm going to believe her. If she tells me that she saw a squirrel steal a police officer's handgun and rob a bank, I'm going to require further evidence.

These are elementary points about the epistemology of testimony. But they allow us to see how peculiar the transmission of fake news is. I'm now going to argue that social media transmission of fake news *is* a form of testimony, but it's a *bent* form of testimony.

Why is it a form of testimony at all? Look at an example. Suppose I believe that Donald Trump threatened to deport Lin-Manuel Miranda (though Miranda is an American citizen). My belief is false; Trump never said that.<sup>8</sup> But I believe it because I saw the headline circulating in my Facebook feed. My belief is therefore held on the basis of testimony; I believe it *because* it was presented as truth by another person.

Yet this is certainly not a standard case of testimony. For one thing, the relationship between the testifier and the content of her testimony is hard to categorize. In standard cases, the testifier makes an *assertion*. But when my friend posts a link to the story about Trump deporting Miranda, without further comment, is my friend *asserting* the content of that story?

Importantly, there is quite a lot of debate about this question, mostly in terms of whether people are rightly held accountable for posting or retweeting defective social media items. For example, in November 2015, Donald Trump himself posted to Twitter an infographic riddled with fake statistics, including the made-up claim that 81% of white homicide victims are killed by African-Americans (the actual figure is 15%) (Greenberg 2015). When challenged by Fox News personality Bill O'Reilly, Trump replied with a defense he has since given for other demonstrably false tweets: "Bill, am I gonna check every statistic? All it was is a retweet. It wasn't from me" (Colvin 2016).

The "just a retweet" defense will be familiar to social media users. When called out for posting material that is false or offensive, people often insist (truthfully) that they are not the originator of the content—they only passed it along. They often insist that "a retweet is not an endorsement" and claim that they pass along content to encourage discussion, not necessarily to stand behind it.

Is a retweet an endorsement? When you post a news link to Facebook without comment, are you vouching for its truth? These are *disputed norms* of communication. Social media is a relatively new way of distributing information, and we have yet to settle on norms for how to interpret its use. We understand that a newspaper article with an embedded quotation isn't necessarily affirming the content of the quote. But we don't yet have a common understanding about social media shares.

Notice that where we do have established norms for older forms of communication, they can be quite nuanced. Consider this situation: a person on the streetcorner is handing out printed copies of a newsletter. Should you understand that this person *believes* most of the factual claims contained in the newsletter? It depends. If this person is taking payment, then probably not; newsagents don't necessarily believe (or even read) much of what they sell. On the other hand, if the newsletter distributor isn't being compensated, then it is reasonable to assume they believe the contents of what they are passing out. Why else would they bother doing so, unless they believe that they are communicating important truths?<sup>9</sup>

But, for now at least, social media sharing operates under unstable norms. People are happy to be understood as asserting the contents of shared news stories that turn out accurate (especially if they ‘scooped’ their friends) but insist that they meant no such assertion when trouble emerges. And, for now, our accountability conventions seem to tolerate this instability; we may roll our eyes at “a retweet is not an endorsement,” but we don’t (yet) place most embarrassed retweeters in the same category as outright liars or bullshitters.

The instability of these norms is one reason that I called social media sharing a bent form of testimony. The epistemic relationship between testifier and testimony is ambiguous, as we haven’t yet settled on a norm whereby sharing entails assertion.<sup>10</sup> Nevertheless, many of us *treat* social media sharing as if it *were* ordinary testimony, at least until something goes wrong. This is why the “a retweet is not an endorsement” mantra causes so much argument; many of us implicitly assume that our social media interlocutors do believe what they share, even though we are vaguely aware they may later disclaim it. This is part of what makes social media testimony aberrant.

There is a second reason that social media sharing of fake news is bent testimony: many fake news stories are ridiculous, seemingly violating a basic content-related norm of responsible testimony-reception, yet people accept the testimony anyway. Recall the earlier point that the reasonableness of default acceptance of testimony requires that we suspend confidence when a piece of testimony is radically at odds with what we know about the world. This norm seems to be routinely violated by social media users, who accept extraordinary stories about political enemies on mere say-so.

For example, consider the fiasco surrounding a Washington, DC, pizza parlor. For several weeks in 2016, social media posts circulated claiming that the restaurant’s basement was being used as a hub for child sexual abuse by a Satanic cult including Hillary Clinton and several of her senior staff. On December 4, 28-year-old Edgar Welch allegedly drove from North Carolina to Washington, with a loaded AR-15 rifle at his side, in order to “self-investigate” the allegations. Welch was reportedly shocked to discover no evidence of child sex trafficking in the pizza parlor, and thankfully was arrested without harming anyone (Goldman 2016).

This story is a truly bizarre thing to believe, *even if you think Clinton and her staff are evil*. If a cabal of extremely powerful individuals wished to conduct vile and criminal activities, why would they choose to do so in

the basement of a pizza parlor? Wouldn't their money and influence give them access to far more secure locations? The story is inconsistent with basic facts about human behavior, even assuming the worst of Clinton. Sensible employment of testimonial norms ought to filter out this sort of ridiculous story.

Yet Welch was not the only one who believed the pizza parlor story. According to a survey by YouGov and *The Economist*, 46% of Trump voters continued to believe the underlying conspiracy theory, even after the media attention focused on Welch's folly (Frankovic 2016).

There is something about social media sharing that seems to deaden people's normal application of consistency-with-the-world filtering on testimony. Before the era of social media, the pizza parlor story might have circulated by word of mouth among a particular paranoid sub-population (just as lurid urban legends have always surrounded the Clintons). But something about Facebook, etc. allowed a ridiculous story to build testimonial momentum to the point of acceptance by more than the furthest fringe.

I suspect that the two bent features of social media testimony are related to one another. Perhaps people are less inclined to subject ridiculous stories to scrutiny *because* we have unstable testimonial norms on social media. A friend posts a ridiculous story, without comment, and *maybe* they don't really mean it. But then other friends 'like' the story, or comment with earnest revulsion, or share it themselves. Each of these individual communicative acts involves some ambiguity in the speaker's testimonial intentions. But, when all appear summed together, this ambiguity seems to wash away. Perhaps the implicit thought is like this: could it really be that *all* these people aren't *really* testifying to this? A thought like that might overwhelm ordinary skepticism about ridiculous testimony.

I am not sure that I've got this mechanism quite right. Clearly, this is an empirical hypothesis for social scientific investigation. But however the details go, it seems plausible that the bent aspects of social media testimony play a role in the transmission of fake news.

#### THE (INDIVIDUAL) EPISTEMIC VIRTUE OF PARTISANSHIP

So far I have tried to explain what fake news is and how it passes through testimony. These have been descriptive analyses. I'll now turn to normative evaluation of social media transmission of fake news. I will defend a surprising conclusion: even though fake news is false and damaging, the testimonial practices propelling it are consistent with



individual epistemic virtue. Specifically, I will argue it is *partisanship* that makes some fake-news-conveying practices reasonable, and partisanship is consistent with epistemic virtue.<sup>11</sup>

Before giving this argument I need to stress what I am *not* claiming. I am *not* claiming that it is good, full stop, for fake news to circulate and affect our collective political choices. That is obviously not a good thing. Nor am I claiming that partisanship is good in itself. What I *am* claiming is something more nuanced: given the realities of human psychology and politics, certain forms of epistemic partisanship are individually reasonable in the world as we actually confront it. This would not be the case in an ideal world, but that is not where we live. In effect, I am defending a form of *non-ideal* political-epistemic theory. Accordingly, I will argue that our normative focus should be on identifying realistic structural changes, rather than specifying idealized individual practice.

My first claim, then, is that partisanship-in-testimony-reception is sometimes compatible with epistemic virtue. That is, sometimes it makes sense to assign greater credibility to a testifier *because* you know you share a political affiliation with her.

The word ‘sometimes’ is important. I doubt that we should *always* assign greater credibility to co-partisans. If two people are testifying to the spectrometer-observed molecular mass of a particular carbon sample, it is probably not reasonable to trust the Democrat over the Republican, or vice versa. Rather, I mean to defend partisan epistemology within specific domains.

Which domains? The domain of politically normative claims, certainly, such as ‘equality of opportunity is more important than equality of outcome.’ I include also many morally normative claims. And I include some claims that are seemingly purely descriptive when these are relevant to political decisions. Most importantly, I include characterological judgments about particular political candidates. I will call all of these *politically relevant claims*. My position, then, is that partisanship in testimony reception is reasonable with regard to politically relevant claims.

Let me start with testimony about obviously normative matters, political and moral. Here I will simply *assume* that it is sometimes good practice to accept normative testimony from others, though some philosophers dispute this.<sup>12</sup> Allowing this assumption, I am now adding the further claim that it is sometimes reasonable to be differentially receptive to normative testimony from others, depending on their partisan affiliation.

Why? Consider what partisan affiliation involves. Though we sometimes bemoan it as mere tribalism, this is an exaggeration. Partisan affiliation reflects a person's *value commitments*. Political parties are partly defined by common views on the normatively appropriate shape of society. When I learn a person's partisan affiliation, I learn something about the political and moral values she endorses.

Of course, it is foolish to take partisan affiliation as signaling a monolithic set of values. Parties have significant normative diversity *within* them, as shown by recent confrontations between Bernie Sanders and establishment Democrats, or Trump and traditional Republicans. But we can still treat partisan affiliation as a reliable indicator of broad categories of values; that's why many people are hesitant to declare their partisan affiliation in conflict-averse social contexts.

So, when I learn that another person shares my partisan affiliation, I learn that she and I share at least some significant number of normative values. Or, to put it another way, I learn that she *tends to get normative questions right* (by my normative lights). She establishes herself as a more reliable normative judge than I would take her to be by default, or especially if she were affiliated to an opposed party.

Another way to make the point is to think about whether another person is my epistemic *peer* in normative domains. Typically, I should accept testimony only from those who are (reasonably assumed to be) at least as good as I am at making judgments in the domain about which they testify. Regarding sensory judgments (e.g., "Look over there, that's Lenny Kravitz!"), part of the motivation for default acceptance of testimony is that we assume others' perceptual systems are similar to ours (Foley 2001). However, if a person makes repeated perceptual errors, then I should cease regarding her as a peer and discount her perceptual testimony.

Presumably something similar applies to normative testimony. If a person repeatedly makes normatively suspect claims, I should begin to doubt that she is my normative peer, and eventually I should discount her normative testimony. Adam Elga suggests that in normative domains, our *only* epistemic peers are those who agree with us on a broad swath of claims. He offers the following case:

[C]onsider Ann and Beth, two friends who stand at opposite ends of the political spectrum. Consider the claim that abortion is morally permissible. Does Ann consider Beth a peer with respect to this claim? That is: setting aside her own reasoning about the abortion claim (and Beth's contrary view about it), does Ann think Beth would be just as likely as her to get things right? (Elga 2007, 492–3)

According to Elga, the answer is no. He reasons that, if Ann and Beth have discussed related issues, then Ann will know that Beth has (according to Ann) many mistaken views about such things as “whether human beings have souls, whether it is permissible to withhold treatment from certain terminally ill infants, and whether rights figure prominently in a correct ethical theory” (Elga 2007, 493). Since Beth is wrong about these issues, it would be reasonable for Ann to treat her moral testimony about abortion as less than that of a peer. On the other hand, Elga says, if they agreed about these other issues, then Ann could reasonably regard Beth as a peer on the abortion question. The upshot, says Elga, is that “with respect to many controversial issues, the associates who one counts as peers tend to have views that are similar to one’s own” (Elga 2007, 494).<sup>13</sup>

We can infer many of a person’s normative beliefs from her partisan affiliation, so partisan affiliation is a reasonable proxy for epistemic peerhood in political and moral normative domains. And social media participants tend to group themselves into partisan networks (Bakshy et al. 2015). People often know the partisan affiliation of their social media contacts, especially those who regularly post links to political news. Social media users treat these partisan signals as indicators of whom they can regard as normative peers, and this allows them to decide which testimony to receive.

So far I have been talking about testimony that is overtly normative—claims about how we *should* live together. But much of fake news is ostensibly descriptive. It claims that such-and-such happened, or that so-and-so said something. These are not normative claims. Is it reasonable to use partisan affiliation, which I’ve claimed is an indicator of *normative* peerhood, to assess testimony about descriptive claims?

I think so, at least when the testimony is politically related. This is because the act of transmitting political news implicates *normative decisions* on the part of the testifier. Often these are decisions about what is politically *important*. Our audience has only so much time to spend reading about politics, so we need to avoid wasting this time on trivialities. Political importance is a value-laden notion; the set of topics that are important to a political conservative will not be identical with those important to a progressive. Importance also plays a role in weighing the degree of confidence one must have before relaying an uncertain news report. Typically, the need to be confident scales with the importance of the subject matter (though this can be complicated in cases that require urgent response).

My point: trusting a testifier regarding political news requires trusting her judgments about political importance, and this means believing that she has, by and large, the right political values. Partisan affiliation provides this information.

This is especially true when testifiers offer *characterological evidence* about particular political candidates. Characterological evidence is dangerous because it can be easily biased through selective reporting. Everyone (political candidate or not) has some negative traits and has made some bad choices. If I choose to report only negative information, deliberately excluding redeeming characteristics, then relying on my testimony could lead you to a harsh assessment of *anyone's* character. Hence, if you are going to rely on my testimony about events related to a candidate's character, you need to trust that I have good judgment about the representativeness of particular stories, whether additional details might be exculpatory, and so forth. You need to trust that you and I share values relevant to these judgments—and partisan affiliation is a good indicator.

I've argued, then, that partisanship can be relevant to assessing the trustworthiness of testifiers on politically relevant claims—not just openly normative claims, but also some related descriptive claims about events, especially those that are meant to provide characterological evidence about candidates. If I am right, then some degree of partisanship-in-testimony-reception is indeed compatible with epistemic virtue.

Of course, the 'some degree' qualifier is important. Many epistemic virtues can become vicious in excess. Skepticism is "healthy" in moderation, but becomes destructive as it grows. One form of epistemic injustice is 'credibility excess,' in which we grant inappropriately high testimonial credibility on the basis of a testifier's demography.<sup>14</sup> These are types of epistemic vice that come from overextending virtuous practice. Similarly, it is obviously possible to make an epistemic vice of partisanship. One can overextend the credibility granted to co-partisans, either by simply assigning too much credibility or by allowing it to intrude into non-politically-relevant domains.<sup>15</sup>

Hence, my claim is obviously not that one should always believe whatever testimony is given by one's co-partisans. That is false. But, generally speaking, one *may* (and perhaps should) attribute greater credibility to co-partisan testifiers than to others. This is simply reasonable, given that shared partisan affiliation points to shared normative values.

And this, finally, allows us to see why it is individually reasonable to accept the bent testimony of social media sharing. Recall the core

ambiguity of bent testimony: people who share stories on social media may or may not be lending their epistemic imprimatur (“a retweet is not an endorsement”—unless it is), yet we tend to treat what they share as if it were unambiguous testimony anyway. I suggest that we do this *because* the reasonable credibility boost we give to co-partisans overcomes the hesitancy we feel about bent testimony.

The model is like this: I read a story on social media, shared by one or two of my co-partisan friends. The story is shocking, and I am vaguely aware that my friends’ communicative intentions are ambiguous. Maybe they aren’t really putting their imprimatur on this story. But I know that these friends share my partisan affiliation, hence many of my normative values. They wouldn’t lie to me, right? They would exercise reasonable judgment about balancing confidence in important information, right? They wouldn’t be confused about the relevance of this information to assessing a candidate’s character, right?

Not always right, of course. But right often enough that trusting my co-partisans is reasonable. Hence, despite some qualms over the bent ambiguity of their testimony, I find myself starting to believe the stories they transmit.

Of course, if we were epistemic angels, we’d be more careful to check our testifiers’ sources, to look for independent verification, to ask questions. But all that is true about accepting *any* testimony, not just on social media or among partisans. We take others’ words for it when we just don’t have the time to go out and investigate claims for ourselves. Social media sharing is the same. There is so much information available, and only so much time to conduct inquiries. In an epistemically non-ideal world, given our temporal and cognitive limitations, it simply makes sense to trust others, even when we antecedently know that this will sometimes lead us astray.<sup>16</sup>

Fake news, then, is a bad side effect of an individually reasonable epistemic practice. If we want to solve the problem of fake news, we’re unlikely to find it in demanding revision to individual epistemic choices. Yes, we *could* insist that everyone become a far savvier user of social media testimony. But most people won’t, and in our epistemically non-ideal world, most are reasonable not to bother.

If we want to solve the problem of fake news, we need to look beyond individual epistemic practices—we need to look at institutions.

I've argued that fake news is transmitted through a bent form of testimony and benefits from credibility gained through partisan affiliation. I've also argued that the problem is not likely to be addressed by focusing on partisanship, which had seemed the likely target. Instead, I'll now suggest, the best place to focus is on institutional arrangements that reduce the bentness of social media testimony.

Recall the core bent property of social media testimony: we have an unstable set of norms for assigning testimonial intentions to social media shares. We tend to treat them as conveying our interlocutors' testimonial approval, yet we also sometimes accept that "a retweet is not an endorsement."

It's this ambiguity that allows fake news to slip through. Resolving our ambiguous norms would greatly reduce the effectiveness of fake news. If we firmly established the norm that social media sharers *are* understood as conveying testimonial endorsement, then people would be less likely to share unverified stories, to avoid later being held responsible for errors. Alternatively, if we firmly established the norm that social media shares (without further comment) communicate no testimonial endorsement whatsoever, then people would be less likely to come to believe fake news on the basis of their friends' transmissions.

Which of these norms should we aim to establish? I think we must disqualify the second option on grounds of irreality; it is very unlikely that we will be able to convince people to begin treating social media sharing as communicating no testimonial endorsement whatsoever. After all, people use social media to communicate facts about themselves and their friends: so-and-so had a baby; my partner got a new job; look at these amazing photos of this place I could afford to travel to! It is reasonable to expect these reports to be truthful (allowing for self-promotional burnishing). A norm that required us to selectively withhold trust from a subset of ostensibly factual stories (those that are politically related) transmitted via a medium we usually trust seems unlikely to be psychologically efficacious.

A norm of accountability seems preferable then; we should aim for a norm that denies "a retweet is not an endorsement." People who share news should be unambiguously understood to lend their testimonial endorsement (barring explicit disclaimer), and should be held accountable if their claims are later shown false, in just the same way that a person spreading false rumors about an acquaintance may be held accountable. 'Holding accountable,' of course, needn't involve punishment or even

condemnation. It may be simply a loss of testimonial reputation, such that repeated offenses lead to one's justified discredit as a participant in testimonial exchange.

How do we replace our ambiguous social media testimonial norms with a clear norm of accountability? Unfortunately, this may not be easily accomplished by individuals. It would require that we each keep track of where we learned every piece of purported political news, track later revelations about their accuracy, and trace debunked stories back to specific social media interlocutors. Sometimes days or weeks pass between the initial promotion of fake news and later debunking. It's not likely at all that many of us could put in the cognitive effort needed to sustain this norm.

This is why institutions matter. When the sustenance of a norm demands unrealistic resources from individual adherents, we can offload these demands to institutions. For a simple example, consider how pedestrians and cyclists interact on busy pathways. Generally, it is best for the pathway to be split—but which side is for pedestrians and which for cyclists? This can vary from place to place, and we cannot assume everyone will remember which is the pedestrian side in every place they visit. A simple institutional solution offloads the demand from memory: paint a line down the middle of the pathway and draw pedestrian and cycle icons on the appropriate sides. The paint facilitates adherence to the norm and makes accountability unambiguous.

We need something similar for social media testimony. The obvious source of infrastructure is the social media platforms themselves. If we could offload onto them the demands of keeping track of who-testified-to-what, then we could sustain a norm of holding people accountable for sharing fake news.

This may sound worrisomely as if I am calling for social media platforms to arbitrate the veracity of news stories and then censor their own users. Social media platforms will certainly refuse to do these things, and we probably would not want them to anyway. But there are milder solutions. I will conclude by describing one—though probably a better solution can be devised by professional technology and communication specialists.

We can start with something social media platforms are *already* doing. On December 16, 2016, Facebook announced that it will implement new measures against the distribution of fake news.<sup>17</sup> Users will be able to report stories they believe are false. Facebook will not determine veracity itself; instead it will refer frequently reported stories to independent fact-checking organizations such as Snopes.com. If these organizations judge



a story false, Facebook's system will flag it as 'disputed,' and this flag will be visible to viewers, along with a link to the fact-checker's debunking. Anyone subsequently attempting to share the story will be confronted with a prompt informing them of its disputed status. They can still choose to share the story, but it will be auto-flagged as disputed, and Facebook may weight its algorithm to display other posts ahead of disputed stories.

This set of measures is a good idea, and will certainly help to sustain an accountability norm. A story that has been flagged as disputed is, presumably, less likely to be trusted on the basis of testimony, and people who persist in sharing disputed stories may suffer reputational consequences. But there are limitations to these measures. Most importantly, they may move too slowly. Many social media stories are ephemeral; everyone is talking about the latest outrage today, but by tomorrow they have moved on to the next (especially amid the perpetual chaos of the Trump administration). Facebook's reporting-and-referral-to-Snopes method will take time to catch up with individual fake news stories. By the time a story has been flagged disputed, much of the audience will have already seen it. Of course, it is good to take measures that reduce the durability of fake news by warning latecomers. But it would be better still to get ahead of the next fake story.

Hence, I suggest that social media platforms provide the infrastructure for tracking the testimonial reputation of individual *users*. Facebook already knows exactly what each user chooses to share. It will also soon have a database of disputed stories, courtesy of the measures it began implementing in December. It would be computationally simple, then, for Facebook to calculate a Reputation Score for individual users, based upon the frequency with which each user chose to share disputed stories. Reputation Scores could be displayed in a subtle way, perhaps with a colored icon beside user photos.<sup>18</sup>

Note that this proposal does not involve censorship. Facebook would not prevent anyone from sharing or receiving any story. Individual users could choose to ignore Reputation Scores. Facebook could provide optional settings for users' News Feeds, allowing them to deprioritize posts from those with low Reputation Scores, but this need not be the default.

The key advantage of this system is that it offloads memory resources for testimonial track records from individual users to institutional infrastructure. Doing so would encourage a norm of accountability for social media sharing; people could easily identify those who routinely share debunked stories, and the "retweet is not an endorsement" line would



become increasingly implausible as track records tabulated. Gradually, our ambiguous testimonial norm could be displaced by a norm expecting genuine endorsement from unadorned sharing.

There are some problems with this proposal, of course. One is the danger of diluting the viewpoint neutrality of social media platforms. Some people continue to insist that fake news stories are true, even after repeated debunking. Months after the pizza parlor debacle, it is still possible to find new blog posts exploring the “conspiracy” and denouncing social media platforms for “censoring” discussion. My proposal would surely result in such people being assigned very unfavorable Reputation Scores by Facebook, and presumably the platform would prefer not to alienate any users. But I think that this problem is unavoidable for any serious institutional response to fake news. Notice that Facebook’s new measures already depart from viewpoint neutrality by tagging stories as disputed.

A more particular problem with my proposal is that it sets a worrisome precedent for social media platforms ‘ranking’ their users. Dystopic speculative fiction regularly imagines that we will spend much of our future struggling to secure positive ratings for our social media personae—see Gary Shteyngart’s novel *Super Sad True Love Story*, or the *Black Mirror* episode “Nosedive.” This anxiety becomes especially pressing with repressive governments; recently, some local authorities in China began calculating a ‘social-credit score’ for citizens, which determines access to some government services and allegedly may include politically related social media behavior (*The Economist* 2016). Perhaps the danger of such possibilities is severe enough that we should avoid any possible precedent, including the Reputation Score of my proposal.

I am sure there is a better, subtler, solution than my specific proposal. My fundamental point is only that we should start thinking in institutional terms. We need to resolve the ambiguous norms that make social media testimony so bent. Proposals addressing partisanship or other aspects of individual epistemic virtues are unlikely to work—partly because, as I’ve argued, some partisanship in testimony is individually reasonable. The most plausible solutions will be institutional, and social media platforms must do *something* to provide infrastructure for an accountability norm. Better norms, facilitated by wise institutions, are what will stop fake news exploiting gaps in otherwise reasonable norms of communication and belief.

ACKNOWLEDGMENTS

Thanks to the editor and an anonymous referee for helpful suggestions on this paper.

NOTES

1. For the Clinton story, see Snopes.com (2016a). For the Pence story, see Snopes.com (2016b).
2. Silverman (2016) claims that the top 20 fake news headlines of the 2016 election cycle generated more Facebook engagement than the top 20 headlines from reputable news sources. Allcott and Gentzkow (2017), however, argue that fake news was unlikely to have been the determining factor in Trump's victory.
3. Addendum, July 2017: This paper was written in January and February 2017. Since that time, the term 'fake news' has acquired an additional use, especially in tweets by President Trump. In this new usage, 'fake news' seems to mean any form of reportage that the speaker disagrees with. For example, on February 6 2017 President Trump tweeted: "Any negative polls are fake news, just like the CNN, ABC, NBC polls in the election. Sorry, people want border security and extreme vetting" (<https://twitter.com/realdonaldtrump/status/828574430800539648>). A further innovation is the extension of the term from particular news *stories* to entire news organizations. We can see this evolution in Trump tweets like "Mainstream (FAKE) media refuses to state our long list of achievements, including 28 legislative signings, strong borders & great optimism!" (April 29 2017, <https://twitter.com/realdonaldtrump/status/858375278686613504>) and "The Fake News Media works hard at disparaging & demeaning my use of social media because they don't want America to hear the real story!" (May 28 2017, <https://twitter.com/realdonaldtrump/status/868985285207629825>). However, this seems to be an idiosyncratic use of the term among Trump and his affiliates. *This paper will persist with analysis of the original use of 'fake news', as it emerged during the 2016 campaign.*
4. In November 2016, the *Washington Post* described claims by anonymous "experts" at the website PropOrNot that many fabricated anti-Clinton stories were amplified by Russian propaganda organs (Timberg 2016). But other journalists dispute the reliability of these claims (Norton and Greenwald 2016).
5. Note that this definition excludes satirical news of the sort featured in *The Daily Show* or *The Onion*. Satire does not typically aim to deceive; its comedic effect relies upon the audience appreciating that it is engaged in exaggeration

or parody. There is, however, a peculiar genre of Internet pseudo-satire, one that is carefully designed to trick *some but not all* readers. The joke is on gullible people, who are meant to earnestly share dross and then be snickered at by their savvier social media friends. The goal of pseudo-satire, splitting the audience into savvy jokers and gullible butts-of-jokes, distinguishes it from typical fake news. Typical fake news does not *require* that any of the audience see through the deception, and of course is usually intended to be believed by as many as possible.

6. A bullshitter, in the sense identified by Harry Frankfurt (2005), is distinct from a liar. A liar makes claims she knows to be false. A bullshitter makes claims that may or may not be true; she is indifferent to whether they turn out right, though she wants others to believe her regardless. Some creators of fake news, especially those with a commercial motive, are technically bullshitters rather than liars.
7. Our existing practices for filtering testifiers are defective in a number of ways. One important way is that we tend to allow a person's apparent race or gender to affect the degree of credibility we assign them. This is epistemically non-ideal and a form of injustice (Fricker 2007).
8. Snopes.com (2016c).
9. Of course, this does not mean that *you* should believe the contents of their newsletter. A reliable imperative of city living is to avoid accepting any piece of paper handed out on streetcorners.
10. Technically, a communicative act isn't testimony at all if the 'speaker' does not intend to imply the truthfulness of what they communicate. So, if "a retweet is not an endorsement" is right, then purportedly factual retweets and shares cannot be testimony. But I will stick with talking about 'testimony,' since we don't yet have another word for ambiguous speech acts that may or may not be testimony depending on as-yet-unsettled communicative norms.
11. A brief autobiographical digression: I imagine some readers will assume my position is motivated by personal inclination. They will assume that I am a dedicated partisan seeking to vindicate my own opposition-flaying practices. But the truth is the opposite. My own inclinations are anti-partisan; I tend to irritate comrades by policing their insufficient interpretive charity toward our opponents. I am perhaps less inclined than most to engage in partisan epistemic filtering. (One should, of course, be extremely cautious about introspectively attributing exceptional epistemic practices to oneself (Kruger and Dunning 1999; Pronin et al. 2002). What I claim here is based upon others' frustrated descriptions of my disappointingly unpartisan responses.) In fact, I argue elsewhere that we have strong moral duties to aim to understand, and

even empathize with, those with whom we disagree on moral and political topics. Hence the position I defend here is not a natural one for me to take. My motivation for adopting it is a second-order extension of my commitment to empathizing with those with whom I disagree; here I am trying to empathize with those who disagree with me *about* the practice of political disagreement!

12. See Hills (2009) for challenges to moral testimony, and Sliwa (2012) for a defense.
13. Sarah McGrath (2007) challenges Elga on this point: she notes that even if Ann and Beth disagree on issues adjacent to abortion, they probably agree on many background moral beliefs about e.g., lying, murder, slavery, etc. Given the great frequency with which they *do* agree, Ann should regard Beth as a peer after all. For my part, I'm not sure. It's not clear that we have guidelines for which or how many topics *should* count as "related" when we assess peerhood with respect to a particular judgment. I can't settle that here.
14. The term 'credibility excess' comes from Fricker (2007), though Fricker herself argues that epistemic injustice is primarily a problem of credibility deficit. But see Medina (2011) and Davis (2016) for different views.
15. At extremes, partisan fragmentation risks undermining the sharing of normative reasons that is essential to democratic citizenship. See Lynch (2016, chapter 3) for worries about the Internet's role in accelerating this trend.
16. I am not denying that individual people *can* improve their personal practices for using social media. One easy improvement is to discredit links to news sources with a history of misleading or false reporting. Two prominent lists of these sources are available from Zimdars (2016) and Brayton (2016).
17. See Mosseri (2016).
18. A complication: what if I want to share a story that I know is false, precisely in order to explicitly alert my audience to its falseness? How would the Reputation Score algorithm avoid counting this against me? One solution might be to allow me to attach a disputed flag to a link *myself* as I post it, thus explicitly signaling that I am *not* endorsing the story. Such pre-tagged disputed stories would not count against one's Reputation Score.

## REFERENCES

- Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31 (2): 211–236.
- Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science* 348 (6239): 1130–1132.

- Brayton, Ed. 2016. "Please Stop Sharing Links to These Sites." *Patheos*, September 18. <http://www.patheos.com/blogs/dispatches/2016/09/18/please-stop-sharing-links-to-these-sites/>.
- Coady, C. A. J. 1992. *Testimony: A Philosophical Study*. Oxford University Press.
- Colvin, Jill. 2016. "Donald Trump's Favorite Excuse? It Was Just a 'Retweet'." *Associated Press*, February 22. <http://elections.ap.org/content/donald-trumps-favorite-excuse-it-was-just-retweet>.
- Davis, Emmalon. 2016. "Typecasts, Tokens, and Spokespersons: A Case for Credibility Excess as Testimonial Injustice." *Hypatia* 31 (3): 485–501.
- The Economist*. 2016. "China Invents the Digital Totalitarian State." December 16. <http://www.economist.com/news/briefing/21711902-worrying-implications-its-social-credit-project-china-invents-digital-totalitarian>.
- Elga, Adam. 2007. "Reflection and Disagreement." *Nous* 41 (3): 478–502.
- Foley, Richard. 2001. *Intellectual Trust in Oneself and Others*. Cambridge: Cambridge University Press.
- Frankfurt, Harry. 2005. *On Bullshit*. Princeton University Press.
- Frankovic, Kathy. 2016. "Belief in Conspiracies Largely Depends on Political Identity." *YouGov*, December 17–20. <https://today.yougov.com/news/2016/12/27/belief-conspiracies-largely-depends-political-iden/>.
- Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.
- Gibbs, Nancy. 2017. "A Note to Our Readers." *TIME*. January 24. <http://time.com/4645541/donald-trump-white-house-oval-office/>.
- Goldberg, Sanford. 2010. *Relying on Others: An Essay in Epistemology*. Oxford University Press.
- Goldman, Adam. 2016. "The Comet Ping Pong Gunman Answers Our Reporter's Questions." *The New York Times*, December 7. <http://www.nytimes.com/2016/12/07/us/edgar-welch-comet-pizza-fake-news.html>.
- Greenberg, Jon. 2015. "Trump's Pants on Fire Tweet that Blacks Killed 81% of White Homicide Victims." *Politifact*, November 23. <http://www.politifact.com/truth-o-meter/statements/2015/nov/23/donald-trump/trump-tweet-blacks-white-homicide-victims/>.
- Hills, Alison. 2009. "Moral Testimony and Moral Epistemology." *Ethics* 120 (1): 94–127.
- Kruger, Justin, and David Dunning. 1999. "Unskilled and Unaware of it: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments." *Journal of Personality and Social Psychology* 77 (6): 1121–34.
- Lackey, Jennifer. 2008. *Learning from Words: Testimony as a Source of Knowledge*. Oxford University Press.

- Lynch, Michael Patrick. 2016. *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data*. New York: Liveright Publishing.
- McGrath, Sarah. 2007. "Moral Disagreement and Moral Expertise." In *Oxford Studies in Metaethics Vol. 4*, edited by Russ Shafer-Landau, 87–108. Oxford University Press.
- Medina, José. 2011. "The Relevance of Credibility Excess in a Proportional View of Epistemic Injustice: Differential Epistemic Authority and the Social Imaginary." *Social Epistemology* 25 (1): 15–35.
- Mosseri, Adam. 2016. "News Feed FYI: Addressing Hoaxes and Fake News." *Facebook Newsroom*, December 15. <http://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/>.
- Norton, Ben, and Glenn Greenwald. 2016. "Washington Post Disgracefully Promotes a McCarthyite Blacklist From a New, Hidden, and Very Shady Group." *The Intercept*, November 26. <https://theintercept.com/2016/11/26/washington-post-disgracefully-promotes-a-mccarthyite-blacklist-from-a-new-hidden-and-very-shady-group/>.
- Pronin, Emily, Daniel Y. Lin, and Lee Ross. 2002. "The Bias Blind Spot: Perceptions of Bias in Self Versus Others." *Personality and Social Psychology Bulletin* 28 (3): 369–381.
- Rainey, James. 2017. "20th Century Fox Apologizes for 'A Cure for Wellness' Fake News Promos." *Variety*, February 16. <http://variety.com/2017/film/news/cure-for-wellness-fake-news-promos-studio-apologizes-1201990634/>.
- Silverman, Craig. 2016. "This Analysis Shows How Viral Fake News Stories Outperformed Real News on Facebook." *BuzzFeed*, November 16. <https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook>.
- Silverman, Craig, and Lawrence Alexander. 2016. "How Teens in The Balkans Are Duping Trump Supporters With Fake News." *BuzzFeed*, November 3. <https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo>.
- Sliwa, Paulina. 2012. "In Defense of Moral Testimony." *Philosophical Studies* 158 (2): 175–195.
- Snopes.com. 2016a. "Shots Hired." October 14. <http://www.snopes.com/wikileaks-cofrms-hillary-clinton-sold-weapons-to-isis/>.
- Snopes.com. 2016b. "A 'Vulgar' Lie." October 15. <http://www.snopes.com/mike-pence-calls-michelle-obama-vulgar/>.
- Snopes.com. 2016c. "Puerto Rico Suave." November 4. <http://www.snopes.com/trump-to-deport-lin-manuel-miranda/>.

- Timberg, Craig. 2016. "Russian Propaganda Effort Helped Spread 'Fake News' During Election, Experts Say." *The Washington Post*, November 24. [https://www.washingtonpost.com/business/economy/russian-propaganda-effort-helped-spread-fake-news-during-election-experts-say/2016/11/24/793903b6-8a40-4ca9-b712-716af66098fe\\_story.html](https://www.washingtonpost.com/business/economy/russian-propaganda-effort-helped-spread-fake-news-during-election-experts-say/2016/11/24/793903b6-8a40-4ca9-b712-716af66098fe_story.html).
- Zimdars, Melissa. 2016. "False, Misleading, Clickbait-y, and Satirical 'News' Sources." <http://d279m997dpfwgl.cloudfront.net/wp/2016/11/Resource-False-Misleading-Clickbait-y-and-Satirical-%E2%80%9CNews%E2%80%9D-Sources-1.pdf>.