

## Commentaries

# A Psychological Approach to Promoting Truth in Politics: The Pro-Truth Pledge

Gleb Tsipursky<sup>\*a</sup>, Fabio Votta<sup>b</sup>, James A. Mulick<sup>c</sup>

[a] Decision Sciences Collaborative, The Ohio State University, Columbus, OH, USA. [b] Department of Social Sciences, University of Stuttgart, Stuttgart, Germany. [c] Department of Psychology, The Ohio State University, Columbus, OH, USA.

## Abstract

Some recent psychology research has shown why people engage in deceptive behavior, and how we can prevent them from doing so. Given the alarming amount of fake news in the US public sphere, a group of psychologists has sought to combine the available research in a proposed intervention, the Pro-Truth Pledge, to help address this problem. The pledge asks signees to commit to 12 behaviors that research in psychology shows correlate with an orientation toward truthfulness. Early results show both that private citizens and public figures are willing to take the pledge, and initial survey, interview, and observational evidence shows the effectiveness of the pledge on reducing sharing misinformation on social media.

**Keywords:** deception, post-truth politics, pro-social behavior, fake news, alternative facts

Journal of Social and Political Psychology, 2018, Vol. 6(2), 271–290, doi:10.5964/jspp.v6i2.856

Received: 2017-09-03. Accepted: 2018-02-15. Published (VoR): 2018-07-27.

Handling Editor: Winnifred Louis, University of Queensland, Brisbane, Australia

\*Corresponding author at: Decision Sciences Collaborative, The Ohio State University, 450 Wetmore road, Columbus, OH, 43214, USA. E-mail: gleb@intentionalinsights.org



This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fake news articles, memes, videos, tweets, and other misinformation, a category that has recently been termed “viral deception” by Kathleen Hall Jamieson, director of the Annenberg Public Policy Center, are sweeping social media, shared by ordinary citizens ([The Annenberg Public Policy Center of the University of Pennsylvania, 2017](#)). The impact of such viral deception, together with extensive misinformation spread by high-profile political figures during the 2016 US presidential campaign and the “Vote Leave” campaign in the UK Brexit referendum, have caused the venerable Oxford Dictionary to choose as the 2016 word of the year *post-truth*, “circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief (“[Word of the year 2016 is...](#),” 2016).” Likewise, Collins Dictionary chose “fake news” as its 2017 word of the year, meaning “false, often sensational, information disseminated under the guise of news reporting (“[The Collins Word of the Year 2017 is...](#),” 2017).”

Sharing such misinformation – at least by private citizens – is not necessarily intended to harm others or even deliberately deceive. Our emotions and intuitions focus more on protecting our worldview and personal identity,

and less on finding out the most accurate information (Haidt, 2012; McDermott, 2004; Nyhan & Reifler, 2010). Still, regardless of the intentions, the impact of sharing this misinformation is vast. A study showed that in the three months before the US presidential election, the top 20 fake election-related news stories on Facebook received more engagements – reactions, comments, and shares – than the top 20 real news stories, 8,711,000 compared to 7,367,000 (Silverman, 2016). Thus, real news in this case has been outcompeted by fake news. Another study, which looked at a wider number of fake news stories, showed that in the same period of three months before the 2016 US Presidential election, 156 misleading news stories got just under 38 million shares on Facebook (Allcott & Gentzkow, 2017). Note that the researchers in this study examined shares on Facebook rather than engagements: the latter number would have been much higher. Fake news tends to favor conservative perspectives: the same study showed that stories favorable to Donald Trump were shared 30 million times, while those favorable to Hillary Clinton were shared a total of 7.6 million times.

This sharing aligns with the example set by the candidates. For example, *The Washington Post's* well-respected Fact-Checking Column has compared the two major candidates in the US presidential election in early November 2016, and found that Trump received their worst rating for claims fact-checking 63 percent of the time, while Clinton received the worst rating 14.2 percent of the time. Notably, previously, most candidates received the worst rating between 10 and 20 percent of the time (Cilizza, 2016). Likewise, when caught lying by media fact-checkers, Clinton tends to apologize and back away from lies, while Trump tends to double down and attack the fact-checkers. This may help explain why trust among Republicans in the media has fallen by more than half, from 32 to 14 percent, from September 2015 to September 2016 (Swift, 2016). The specific impact of candidates on their supporters sharing false information may be explained in part by the research on emotional contagion, which shows that followers tend to absorb and emulate the emotions of their leaders (Hatfield, Cacioppo, & Rapson, 1993).

How impactful is such sharing? First, we need to recognize that most US adults get news on social media: 62 percent according to a recent poll by the Pew Research Center (Gottfried & Shearer, 2016). Another recent poll by Ipsos, conducted in late November and early December 2016, showed that American adults are prone to be deceived by fake news headlines. Surveying 3,015 adults, the method of this poll involved showing respondents six election-related headlines, three false and three true, and asking if they recognized the headline. In the case that they did recognize the headline, the respondents were asked to rate the headline as “very accurate,” “somewhat accurate,” “not very accurate,” or “not at all accurate.” Of the people who recognized the fake election-related headlines, approximately 75 percent rated the headlines “very accurate” or “somewhat accurate.” Republicans were slightly more likely to be fooled by fake news, rating fake news headlines they recalled as accurate 84 percent of the time, compared to 71 percent for Democrats (Silverman & Singer-Vine, 2016).

This fake news came from a variety of channels, but a major portion appears to have come from Russia's efforts to use its digital propaganda efforts to influence the US (Kwong, 2017; Shu et al., 2017). Of course, political partisans for either side also created massive amounts of fake news (Green & Issenberg, 2016). So did people trying to make money off spreading fake news (Subramanian, 2017).

The US is far from unique in the impact of fake news. The UK was another target of Russian's digital propaganda efforts, with researchers at the University of Edinburgh finding many hundreds of accounts operating by the Russian Internet Research Agency trying to spread fake news to influence UK politics (Booth, Weaver, Hern, & Walker, 2017). Russia-owned accounts spread misinformation in Spain to stir up the Catalan independence movement

(Palmer, 2017). Russia likewise used misinformation to try to influence the German 2017 elections (Shuster, 2017). The 2017 French elections also drew a great deal of fake news, with a substantial amount coming from Russian-backed accounts (Farand, 2017).

Russia is simply the biggest global player in the spread of fake news: there are many other sources of viral deception. Around the controversial impeachment of president Dilma Rousseff in 2016, three out of the five most shared news stories on Facebook were false, generally originating from domestic Brazilian websites. India's introduction of a 2,000-rupee note in November 2016 led to the widespread and false rumor via the social media messaging app Whatsapp that the bill would have a surveillance chip. In Italy, the anti-establishment Five Star Movement has been implicated in spreading false news about the government's activities. In France, anti-abortion groups set up fake abortion information websites that present themselves as neutral and official sites with free helpline numbers, but which actually promote false anti-abortion propaganda (all the examples in this paragraph come from Connolly et al., 2016). Likewise, a YouGov survey of over 70,000 online news consumers in 36 countries shows that only 24 percent of respondents believe that social media does a good job in separating fact from fiction, and many respondents stated that they believe social media dynamics in their country encourage the spread of fake news (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2017).

The problem of fake news around the world has been foreseen already for several years. For instance, the World Economic Forum, which is a nonprofit institution focusing on improving world welfare, has rated the spread of false information online as one of the 10 biggest global problems in 2013 (WEF, 2013). An important question to answer is whether Americans are more vulnerable to fake news than citizens in other countries. We can with certainty state that Americans are more likely to be exposed to fake news, because of the influential impact of US politics around the world due to America's global influence. Moreover, due to the relative lack of efforts by the US government to fight fake news, its citizens are easier to influence. By contrast, many European Union governments have taken systematic steps to fight viral deception. The EU as a whole has recently launched a major effort to set up a High-Level Expert Group representing academics, online platforms, news media and civil society organizations to address fake news (European Commission, 2017).

While Americans may be more likely to be targeted by fake news, and less protected by government intervention, there is no reason to believe they are inherently more vulnerable to being deceived. If citizens of other countries were not vulnerable to being deceived, after all, then the producers of fake news would be targeting only the US. Indeed, research shows that those outside the US are similarly highly susceptible to believing fake news when exposed to it. For example, a research study on misinformation in the 2017 French election found that exposing voting-age French people to deceptive election-related statements resulted in the study subjects believing Le Pen's falsehoods, while fact-checking improved the likelihood that people believed in the actual facts (Barrera, Guriev, Henry, & Zhuravskaya, 2017). Since it is actively harmful for our global society for people to believe in and spread falsehoods, how can we stop this problem? A recent research article by prominent scholars in the field suggested any effort to address the situation "must involve technological solutions incorporating psychological principles, an interdisciplinary approach that we describe as 'technocognition.'" (Lewandowsky, Ecker, & Cook, 2017). In the meantime, a separate group of psychologists have been thinking along the same lines, and have come up with a proposed techno-cognitive intervention we term the Pro-Truth Pledge (PTP). The pledge asks signees to commit to 12 behaviors that research in psychology shows correlate with an orientation toward truthfulness. Early results show both that private citizens and public figures are willing to take the pledge, and interviews, external observations, and quantitative surveys show evidence of the effectiveness of the pledge.

## Truth and the Tragedy of the Commons

Although our society as a whole loses when deception is rampant in the public sphere, individuals who practice deceptive behaviors often gain for their own agendas. This type of situation is known as a “tragedy of the commons,” following a famous article in *Science* by Garret Hardin (1968). Hardin demonstrated that in areas where a group of people share a common resource without any controls on the use of this resource, each individual may well have a strong interest in taking more of the common resource than is their fair share, leading to individual gain at great cost to the community as a whole. A well-known tragedy of the commons is environmental pollution. We all gain from clean air and water, yet individual polluters, from a game-theoretical perspective, may well gain more – at least in the short and medium term – from polluting our environment (Hanley & Folmer, 1998). Pollution of truth is arguably similarly devastating to the atmosphere of trust in our political environment.

Solving tragedies of the commons requires, according to Hardin, “mutual coercion, mutually agreed upon by the majority of the people affected,” so as to prevent these harmful outcomes where a few gain at the cost of everyone else (Hardin, 1968). The environmental movement presents many examples of successful efforts to addressing the tragedy of the commons in environmental pollution (Ostrom, 2015). Only substantial disincentives for polluting outweigh the benefits of polluting from a game theoretical perspective (Fang-yuan, 2007). Particularly illuminating is a theoretical piece by Mark van Vugt describing the application of psychology research to the tragedy of the commons in the environment. His analysis showed that in addition to mutual coercion by an external party such as the government, the commons can be maintained through a combination of providing credible information, appealing to people’s identities, setting up new or changing existing institutions, and shifting the incentives for participants (Van Vugt, 2009).

The research on successful strategies used by the environmental movement fits well with work on choice architecture and libertarian paternalism. “Libertarian paternalism” refers to an approach to private and public institutions that aims to use findings from psychology about problematic human thinking patterns – cognitive biases – to shape human behavior for social good while also respecting individual freedom of choice (Sunstein & Thaler, 2003; Thaler & Sunstein, 1999, 2003). Choice architecture is the method of choice used by libertarian paternalists, through shaping human choices for the welfare of society as a whole, by setting up default options, anticipating errors, giving clear feedback, creating appropriate incentives, and so on (Johnson et al., 2012; Jolls, Sunstein, & Thaler, 1998; Selinger & Whyte, 2011).

### Intervention to Address Pollution of Truth: The Pro-Truth Pledge

The Pro-Truth Pledge (PTP), created by a team of behavioral scientists coming primarily from a psychology background, is informed by strategies that have proven successful in the environmental movement and combines them with choice architecture. The pledge is not a way for pledge organizers to tell people what is the truth, but to get them to adopt research-informed methods meant to orient toward accurate evaluation of reality. In taking the pledge, signees agree to abide by twelve behaviors, which are intended to counteract a number of cognitive

biases that contribute to people believing in and sharing misinformation, an essential aspect of the psychology research informing the content of the pledge itself. The full pledge reads as follows:

I Pledge My Earnest Efforts To:

**Share** the truth

- Verify: fact-check information to confirm it is true before accepting and sharing it
- Balance: share the whole truth, even if some aspects do not support my opinion
- Cite: share my sources so that others can verify my information
- Clarify: distinguish between my opinion and the facts

**Honor** the truth

- Acknowledge: acknowledge when others share true information, even when we disagree otherwise
- Reevaluate: reevaluate if my information is challenged, retract it if I cannot verify it
- Defend: defend others when they come under attack for sharing true information, even when we disagree otherwise
- Align: align my opinions and my actions with true information

**Encourage** the truth

- Fix: ask people to retract information that reliable sources have disproved even if they are my allies
- Educate: compassionately inform those around me to stop using unreliable sources even if these sources support my opinion
- Defer: recognize the opinions of experts as more likely to be accurate when the facts are disputed
- Celebrate: celebrate those who retract incorrect statements and update their beliefs toward the truth

One of the biases that the pledge aims to address is the confirmation bias, our tendency to search for and accept information that aligns with our current beliefs (Nickerson, 1998). Research shows that one way to address the confirmation bias involves asking people to consider and search for evidence that disproves their initial beliefs, so that they would not violate the pledge by sharing misinformation (Hirt & Markman, 1995; Kray & Galinsky, 2003; Lilienfeld, Ammirati, & Landfield, 2009).

To ensure full clarity on what constitutes violations of the pledge, the pledge spells out what misinformation means from the perspective of the PTP: anything that goes against the truth of reality, such as directly lying, lying by omission, or misrepresenting the truth to suit one's own purposes. While sometimes misinformation is blatant, sometimes it is harder to tell, and for these tough calls, the PTP relies on credible fact-checking organizations – the same ones that Facebook uses for its fact-checking program – and/or the scientific consensus, as recognized by meta-analysis studies and statements from influential scientific organizations. The pledge asks people to take time to verify information before sharing it, by going to reliable fact-checking websites or evaluating the scientific consensus on any given topic. By taking time to verify this information, signees get an opportunity to evaluate the accuracy of their information and change their perspective if they do not find credible evidence supporting that information. This aspect of the pledge aims to address the extensive sharing of fake news, both by private citizens

and by public figures (Allcott & Gentzkow, 2017). It also aims to address the repeated sharing of incorrect information, which produces the illusory truth effect, people's belief that a false statement is true due to multiple exposure to and thus growing comfort with the false statement (Fazio, Brashier, Payne, & Marsh, 2015). Likewise, asking people to pause and verify before sharing information will slow down their responses, which has been correlated with making fewer errors and facilitating analytical thinking to counteract belief bias (Pennycook, Cheyne, Koehler, & Fugelsang, 2013).

In the spirit of anticipating errors, an important aspect of choice architecture, the pledge encourages signees to celebrate both others and themselves for retracting incorrect statements and updating their beliefs toward the truth. We anticipate that another problematic factor might be the in-group bias, which causes people to favor those who they perceive to be part of their own group, and vice versa for those who they perceive as part of their out-group (Mullen, Brown, & Smith, 1992; Verkuyten & Nekuee, 1999). To address the in-group bias, the pledge asks people to defend other people who come under attack for sharing accurate information even if they have different values, and to request that those who share inaccurate information retract it, even if they are their friends and allies. The Dunning-Kruger effect is another cognitive bias where those who have less expertise and skills in any given area have an inflated perception of their abilities, or in other words are ignorant of their own ignorance (Dunning, 2011; Ehrlinger et al., 2008; Kruger & Dunning, 1999; Sheldon, Dunning, & Ames, 2014). To address this problem, the pledge calls on signees to "recognize the opinions of those who have substantially more expertise on a topic than myself as more likely to be accurate in their assessments."

In addition to the cognitive biases that facilitate deception, other studies have emerged on motivators for honesty and dishonesty. If people perceived others around them as behaving dishonestly, they were also more likely to behave dishonestly themselves; in turn, if they behaved dishonestly, they perceived others as more likely to behave dishonestly (Gino, Norton, & Ariely, 2010). These two patterns together, once they start, create a self-reinforcing spiral of deception. For our purposes, the parallel is clear. For instance, consider social media sharing of viral deception. A person who spreads such deceptive content will perceive others around them as more likely to spread viral deception than is actually the case; likewise, if that person sees someone sharing misinformation, they will be more likely to share viral deception themselves, as that person's actions provide him with an implicit permission to do so. Similarly applicable to spreading misinformation online, research shows that people are more likely to lie if they believe it benefits their in-group (Mazar & Ariely, 2006). So if someone sees an article favorable to their political in-group, they would be more likely to share it without doing any fact-checking, even if the article inspires some skepticism, by comparison to a neutral article. Doing such promotion of questionable content favorable to one's in-group both helps people feel like activists for their cause, and signals to others in their social media network an alliance around shared values, gaining them social capital. Moreover, research shows that if an expert corrects erroneous information, people tend to not accept and internalize the correction unless they also trust the expert; in turn, trust alone is enough to sway people to accept corrective information (Guillory & Geraci, 2013). Thus, any proposed solution needs to address the perception of dishonesty by others and oneself, address benefits to one's in-group from dishonesty, and perceptions of trustworthiness.

Fortunately, we also have research on what causes people to avoid dishonest behavior. Two articles show some intriguing findings: reminders about ethical behavior made people less likely to lie; getting people to sign an honor code or other commitment contract to honesty before engaging in tasks involving temptation to lie increased honesty; making standards for truthful behavior clear decreased deception (Mazar, Amir, & Ariely, 2008a, 2008b). Evidence shows that personal experience with an issue such as global climate change helps correct misinformation

about it (Myers et al., 2013). Reminders about the reputation costs of making false statements proved effective for reducing misinformation shared by candidates for office (Nyhan & Reifler, 2015). This finding is particularly salient to the potential impact of the pledge on public figures who have a reputation that might be negatively impacted if they are found to share misinformation, due to the accountability mechanism of pledge-takers monitoring each other, especially public figures who have taken the pledge.

In an interesting parallel to the environmental movement, those who chose to commit to recycling by signing a pledge were more likely to follow their commitments in comparison to those who just agreed to recycle (Katzev & Pardini, 1987). Our likelihood of lying is strongly impacted by our social network, making it especially important to address social norms around deception (Mann, Garcia-Rada, Houser, & Ariely, 2014). Dan Ariely and Simon Jones summarize and synthesize the research on what moves us to lie and vice versa in their *The Honest Truth About Dishonesty: How We Lie to Everyone—Especially Ourselves*. In a nutshell, they find that what determines whether people lie or not is not some rational cost-benefit analysis, but a wide variety of seemingly-irrational psychological factors. Crucially, our behavior around deception ties strongly to self-identity and group belonging. People generally wish to maintain a self-identity as essentially truthful and to act within accepted group norms, and so inducing a greater orientation toward the truth requires integrating truth-oriented behaviors into one's identity and group affiliation (Ariely & Jones, 2012). The more of these factors a solution can address, the better.

## Motivations for Taking the Pro-Truth Pledge

Why would people take the pledge? Many are frustrated and disheartened by the prevalence of deception in our society, and especially in our political system. Signing the pledge gives them an opportunity to express their discontent and help move our society toward greater honesty. This type of pro-social desire has been found to be a strong motivator in environmental efforts (Van Lange, De Bruin, Otten, & Joireman 1997; Van Vugt & Samuelson, 1999). Moreover, the pledge draws on research showing the importance of incentives. Signing the pledge gives any individual who signs it greater credibility among their peers who know they signed it, improving their reputation. Individuals who sign the pledge then share about it on their social media and personal networks, put an indicator on their Facebook and Twitter and other social media profile photos showing they signed the pledge, and also put a badge on their websites. They can also put the pledge logo on the back of their business cards. Such reputational rewards have been shown to be vital in addressing tragedies of the commons in the environmental movement (Milinski, Semmann, & Krambeck, 2002). Another reward is getting access to unique resources available to signees, such as a search engine composed of credible sources verified as reliable by the PTP organizers (Pro-Truth Pledge, 2017a). They also get to join a variety of closed communities both online and in their local area available only to pledge signees, where they can rely on the credibility of the information being shared by those who signed the pledge and also support and encourage each other in practicing behaviors advocated by the pledge. We know that peer support has proven helpful in maintaining desired behavior change in contexts such as health behaviors, and we anticipate that such support will help maintain truth-oriented behavior (Westman, Eden, & Shirom, 1985; Zimmerman & Connor, 1989). The pledge appeals to people's identities by asking for those who self-identify as truthful and honest to take the pledge and join the community of pledge-takers. This appeal to identity is informed by psychology research on the environmental movement showed that people who report self-identification with a community tend to engage in behaviors condoned by that community (Van Vugt, 2001).

However, would pledge-takers who are private citizens, and thus have no external monitoring, follow such behaviors upon taking the pledge? Psychology research on precommitment suggests that those who commit to a certain behavioral norm will be more likely to follow it (Ariely & Wertenbroch, 2002). Another factor at play is post-factum justification or choice-supportive bias, where our minds want to perceive our past decisions in a positive light, making us more likely to stick to past commitments (Correia & Festinger, 2014). A related phenomenon is a preference for consistency, which recent research suggests influences many people to make decisions that are consistent with their past decisions (Guadagno & Cialdini, 2010). All these mechanisms would also increase the trust of pledge-takers for other pledge-takers and their fact-checking and corrective activities, and research shows such trust is vital for accepting corrective information (Guillory & Geraci, 2013). Intriguingly, other efforts to promote truthfulness using commitment mechanisms are arising as well recently, such as the Peer Reviewers' Openness Initiative to encourage peer reviewers to "make open practices a pre-condition for more comprehensive review" (Peer Reviewers' Openness Initiative, 2017).

Most relevantly for the PTP, at schools that have honor codes students tend to engage in less academic dishonesty (McCabe & Trevino, 1993; McCabe, Trevino, & Butterfield, 1999). Likewise, signing an honor code before a test tends to decrease cheating compared to signing an honor code at the end of a test (Shu et al., 2012). This evidence is further supported by research from the environmental movement on recycling, which shows that those who chose to commit to recycling by signing a pledge were likely to follow on their commitments in comparison to those who just agreed to recycle (Katzev & Pardini, 1987). By analogy, we hypothesize that taking the PTP will decrease sharing misinformation by shifting the underlying mental habits of thought and feeling that contribute to deceptive behaviors, especially since we are concerned with people not sharing misinformation after they sign the pledge rather than before it (Ariely & Jones, 2012; Frijda, Manstead, & Bem, 2000). Likewise, being informed about what it means to be truthful through the twelve behaviors outlined in the pledge will also likely contribute to avoiding sharing misinformation. This stems from research showing that labels on household appliances that list comparisons of energy use and emissions most effectively change behavior when consumers are already concerned with the environment but lack technical knowledge about the appliances (Dietz, Ostrom, & Stern 2003). The other research finding supporting this comes from studies showing that making standards for truthful behavior clear decreased deception (Mazar, Amir, & Ariely, 2008a, 2008b).

Further strengthening precommitment, post-factum justification, and preference for consistency, pledge-takers have an opportunity to participate in PTP community-oriented activities described above, to sign up for email updates, to have themselves listed in a public database of people who signed the pledge, and to share publicly about taking the pledge. They can also sign up to be a PTP advocate, which consists of any of the following: 1) Promoting the PTP pledge to other private citizens; 2) Advocating for public figures to take the pledge; 3) Monitoring and evaluating whether the public figures stick to their commitment. In the initial sign-ups, about 85 percent signed up for email updates or action alerts, about 50 percent wanted to be listed in a public database, and about 30 percent indicated an interest in being a PTP advocate (we do not have sufficient data on community engagement).

We hypothesize that each of the four distinct activities listed above would make it more likely for people to abide by the tenets of the PTP, based on research from successful environmental movement strategies. We suspect that for those who sign the PTP without signing up for email notifications or other forms of active engagement will have a small or perhaps negligible long-term impact on their behaviors, due to the PTP fading from their mind. After all, research on health behaviors shows that intentions to change behavior often fail before temptations or lack of energy, which in the PTP context we can compare to failing to fact-check an article before sharing it

(Schwarzer, 2008). Still, given that people who have committed to recycling by signing a pledge did practice recycling at a higher rate than those who did not, we may indeed witness some impact. Other research on recycling shows that having information about conservation made people more likely to engage in recycling (Oskamp et al., 1991). Getting email updates about the PTP would serve that function. Studies on recycling also show that getting specific recycling opportunities increased the likelihood of recycling, and the action alerts fill that function for the PTP (Vining & Ebreo, 1992). Knowing that one is being monitored for recycling and may get negative messages if one does not recycle has been shown to increase recycling behavior (Lord, 1994). The parallel for the PTP is choosing to list oneself in a public PTP database and thus make oneself available for monitoring, as well as sharing with one's social network and on social media that one took the PTP. Also supportive of the importance of the latter, studies of consumers buying environmentally-friendly products showed that such purchases stemmed in part from the opportunity to signal environmental friendliness to others as a form of status-seeking, and thus sharing about the PTP would similarly signal truth-friendliness (Griskevicius, Tybur, & Van den Bergh, 2010). Active volunteering and community engagement in recycling programs, such as block-leader programs, proved even more effective in increasing recycling behavior (Burn 1991; Hopper & Nielsen, 1991). By analogy, we anticipate that those who engage actively in PTP volunteering and community-oriented activities, online and in-person, will be even more likely to exhibit truth-oriented behaviors. After all, community belonging is crucial for shaping perceptions of self-identity and social norms, which research has found are so important in determining truth-telling behavior.

## Alternatives and Challenges

The current best alternatives to advancing truth in our political system focus on supporting the work of fact-checking organizations. Noble and worthwhile, these much-needed efforts unfortunately do not address the underlying problem of distrust in fact-checking organizations. For instance, according to a September 2016 Rasmussen Reports survey, only 29 percent of all likely voters in the US trust fact-checking of candidates' statements. The political disparity is enormous, and in-line with previous reporting on the partisan divide – 88 percent of Trump supporters do not trust fact-checkers, while 59 percent of Clinton supporters express trust for fact-checkers (Rasmussen Reports, 2016). This distrust for fact-checkers will not be solved by providing more fact-checking or faster, real-time fact-checking. Indeed, research shows that real-time fact-checking may actually make people more resistant to correct information (Garrett & Weeks, 2013). Such distrust can only be addressed by getting citizens to both care more about the truth and by providing credible information about who is truthful.

The PTP aims to solve these problems through appealing to people's identities and getting them more emotionally invested into truth-oriented behavior, while also providing them with information about who are honest public figures. A secondary effect of the PTP may be to help legitimate trustworthy fact-checking organizations. Indeed, research suggests that training in media literacy is likely to reduce perceptions of bias by the media in reporting on controversial news stories, and the behaviors of the PTP are conducive to higher media literacy (Vraga, Tully, & Rojas, 2009).

Of course, the Pro-Truth Pledge may not work despite the problems with the current best alternatives. Virginity pledges have been shown consistently to delay the onset of sexual behavior (Martino et al., 2008). However, other research has shown that STD rates are comparable among those who took a virginity pledge and those who did not, potentially due to lower rates of condom use and testing by those took the pledge (Brückner & Bearman, 2005). Likewise, people might follow the fact-checking guidelines only in regard to articles that they are

already ideologically motivated to discard. In fact, many people who signed the pledge, or prior to signing the pledge, inquired about the PTP guidelines on fact-checking. To address this, the Pro-Truth Pledge website provides specific guidelines on fact-checking, which include guidelines like “Does the article please you? If so, it may be playing to your subconscious biases. Make extra efforts to check if it is true.” (Pro-Truth Pledge, 2017b).

Thus, the PTP may have mixed results in getting people to avoid sharing misinformation. Public figures may become afraid of signing on after a few suffered the reputational damage that comes from being listed as in contempt of the pledge. Likewise, politicians, media venues, and others who benefit from deceiving the voters will likely target the pledge as they see it gain ground. To fend off these attacks, the pledge organizers must work hard to reach across party lines to get diverse public figures from all sides of the political spectrum to commit to the pledge, but this effort may or may not be successful. Another area of attack may be around the definition of misinformation as used by the PTP, for instance regarding potential bias in selecting fact-checking organizations. In part to ameliorate accusations of such bias, the PTP specifically decided to use the same fact-checking organizations as Facebook uses, since Facebook has a huge financial interest in using only the most high-quality fact-checking venues. Moreover, the PTP – unlike fact-checking organizations – only evaluates those who have chosen to sign the pledge; it is an opt-in mechanism, like the Better Business Bureau, as opposed to fact-checkers who fact-check statements that the fact-checking organization finds relevant.

Another finding that might be potentially problematic for the effectiveness of the pledge shows that citizens often use political figures they support as a guide to what they consider true or false, regardless of the facts (Swire, Berinsky, Lewandowsky, & Ecker, 2017). Counteracting this tendency requires that citizens develop trust and invest support into the Pro-Truth Pledge as a guiding mechanism for candidates they consider credible. Indeed, a number of people who have chosen to take the pledge have expressed that they would consider whether a candidate has taken the pledge a strong factor in choosing which candidates to support with their votes, money, and time.

## Pro-Truth Pledge Impact: Case Studies and Preliminary Survey Data

The PTP was launched in March 2017, and by December 20 had over 4700 pledge-takers, including over 300 public figures, including Peter Singer, Steven Pinker, Jonathan Haidt, and many scholars cited in the references of this paper. It also included over 70 public officials. Online and in-person groups dedicated to the PTP have emerged in over 20 US states, and are starting up in other states as well as abroad. The pledge has received positive coverage in prominent venues, such as *The Guardian* and *The Columbus Dispatch* (Enfield, 2017; Smola, 2017).

When asked for why they take the pledge, people generally report a desire to cast a vote against fake news and demonstrate a personal commitment to honesty. Some also discuss the desire to project a reputation as truth-tellers for the sake of gaining greater credibility among those with whom they engage. We have performed some follow-up conversations with pledge-takers to determine whether the pledge impacted their behaviors. A private citizen, US Army veteran John Kirbow, stated how after taking the pledge, he felt “an open commitment to a certain attitude” to “think hard when I want to play an article or statistic which I’m not completely sold on.” He found the

pledge “really does seem to change one’s habits,” helping push him both to correct his own mistakes with an “attitude of humility and skepticism, and of honesty and moral sincerity,” and also to encourage “friends and peers to do so as well.” Christian pastor and community leader Lorenzo Neal described how he “took the Pro-Truth Pledge because I expect our political leaders at every level of government to speak truth and not deliberately spread misinformation to the people they have been elected to serve. Having taken the pledge myself, I put forth the effort to continually gather information validating stories and headlines before sharing them on my social media outlets.”

All others who chose to participate in follow-up conversations shared similar responses. It is important to note that follow-up conversations are limited by two factors: self-selection and self-reporting. After all, the people likely to respond are those who find the pledge beneficial and impactful. To address this concern, we also engaged in external observations of the behaviors of pledge-takers, and have observed instances where the pledge was involved with people retracting statements.

For instance, a candidate for Congress, Michael Smith, took the Pro-Truth Pledge ([GoFundMe.com](https://www.gofundme.com), 2017). He later posted on his Facebook wall a screenshot of a tweet by Donald Trump criticizing minority and disabled children. After being called out on it, he went and searched Trump’s feed. He could not find the original tweet, and while Trump may have deleted that tweet, the candidate edited his own Facebook post to say that “Due to a Truth Pledge I have taken I have to say I have not been able to verify this post (Smith, 2017).” He indicated that he would be more careful with future postings. In another case, Mark Kauffman, a photographer from New York, shared an article from *OccupyDemocrats.com*, a site shown by credible fact-checkers used by the PTP to be systematically unreliable. Other pledge-takers, following the behavior of asking people to stop using unreliable sources regardless of the credibility of the article, asked him to withdraw it, and he did so.

We have also conducted a preliminary survey evaluation of pledge-takers to see whether their sharing of information on social media was impacted by the pledge. We decided to target Facebook, as the most popular social media platform: 44 percent of US adults got news via Facebook in 2016 ([Pew Research Center](https://www.pewresearch.org), 2016).

Our hypothesis was that taking the PTP will impact sharing on Facebook, both when people share news-relevant content themselves on their own Facebook profile, and when they engage in other venues on Facebook, such as Facebook groups or other people’s profiles. Examples of engaging in other venues would include behaviors like asking people to retract incorrect statements, as was the case with pledge-takers asking Michael Smith and Mark Kauffman to retract their statements. To test this hypothesis, we have conducted a study of people who took the PTP and engage actively in sharing news-relevant content on Facebook. Since these people already care about the truth – otherwise, they would presumably not take the pledge – any difference between sharing behaviors before and after taking the pledge can be attributed to finding out about the pledge and taking it.

The full method and results are given in the supplementary materials, and summarized here. We had participants fill out Likert scale (1-5) surveys self-reporting their Facebook engagement with news-relevant content on their own profiles and also with other people’s posts and in groups before and after they took the pledge, with 1 at lowest level of alignment to the pledge behaviors and 5 being full alignment. To avoid the Hawthorne effect of study participants being impacted by observation, the study did not evaluate current behavior, but past behavior. We only recruited participants who took the pledge 4 or more weeks ago to fill out the survey, and asked them about their behavior after taking the pledge. Giving them this period also gave people an opportunity to have the

immediate impact of taking the pledge fade from their mind, thus enabling an evaluation of the medium-term impact of the PTP on sharing news-relevant content.

This study method was informed by the approaches used by studies of whether honor codes address cheating, which is the most comparable form of intervention to the PTP. Such studies similarly rely on self-reporting by students on whether they have cheated or not cheated (McCabe & Trevino, 1993; McCabe, Trevino, & Butterfield, 1999). Similarly, studies of whether virginity pledges delay sexual onset similarly rely on self-reporting (Brückner & Bearman, 2005).

Thus, our method of evaluating the PTP faces the same problems faced by those studies: self-reporting and self-selection. The self-selection impact might be problematic because people self-selected to participate in the study, and those on whom the pledge did not have an impact might choose to avoid participating. The self-reporting might be problematic because people might perceive a behavior change that did not actually occur. For example, they might engage in more fact-checking, but only of articles that did not align with their partisan motivations, despite the clear guidelines on fact-checking on the PTP website.

With these limitations in mind, our early survey results of 24 people who took the pledge suggest that taking the pledge results in a statistically significant increase in alignment with the behaviors of the pledge, both on one's own profile on Facebook and when interacting with other people's posts and in groups. Specifically, on one's own Facebook profile, the median alignment with the PTP score before taking the PTP is 4 ( $SD = 1.14$ ), and the median alignment score after taking the PTP is 4.5 ( $SD = 0.51$ ). For engaging with news-worthy content on other people's profiles and in groups, the median PTP alignment score before taking the Truth Pledge is 3.5 ( $SD = 1.06$ ). The median PTP alignment score after taking the Truth Pledge is 4.5 ( $SD = 0.65$ ). For sharing content on their own profile, 70.83% of participants (17 of 24 respondents) reported an increase of their PTP alignment after taking the PTP, eleven participants increased by one point on the alignment scale, five by two points, one by three points, while the rest maintained the same score. For sharing content in groups and on other people's walls, again, 70.83% of participants reported an increase of their PTP alignment after taking the PTP, twelve participants increased by one point on the alignment scale, five by two, while the rest maintained their initial score. Figure 1 below provides a visual summary of the preliminary survey data.

The results contradict the hypothesis that all those who take the PTP are already honest and that taking the pledge is simply a way to signal their honesty publicly. If that was the case, there would be no statistically significant increase in alignment with the truth-oriented behaviors in the PTP before and after taking the pledge. These results are suggestive but far from conclusive that taking the PTP indeed significantly improves the quality of social media sharing.

These results should not surprise us, since the PTP is a technocognitive intervention based on well-established research on what kind of behaviors are likely to lead to people avoiding deception and sharing facts. The PTP uses all four components shown by psychological research on environmental pollution as crucial to addressing tragedies of the commons (Milinski, Semmann, & Krambeck, 2002; Van Vugt, 2009). It provides information about the credibility of those who sign it, as well as information about what it means to orient toward the truth and what constitutes credible information sources. It appeals to the identity of people to desire to be honest and be perceived that way. Finally, it offers positive reputational rewards for honesty, taking advantage of the psychology research on incentives. In fact, it would be surprising if it did not have a statistically significant impact, at least in the self-selected sample who self-reported their behaviors.

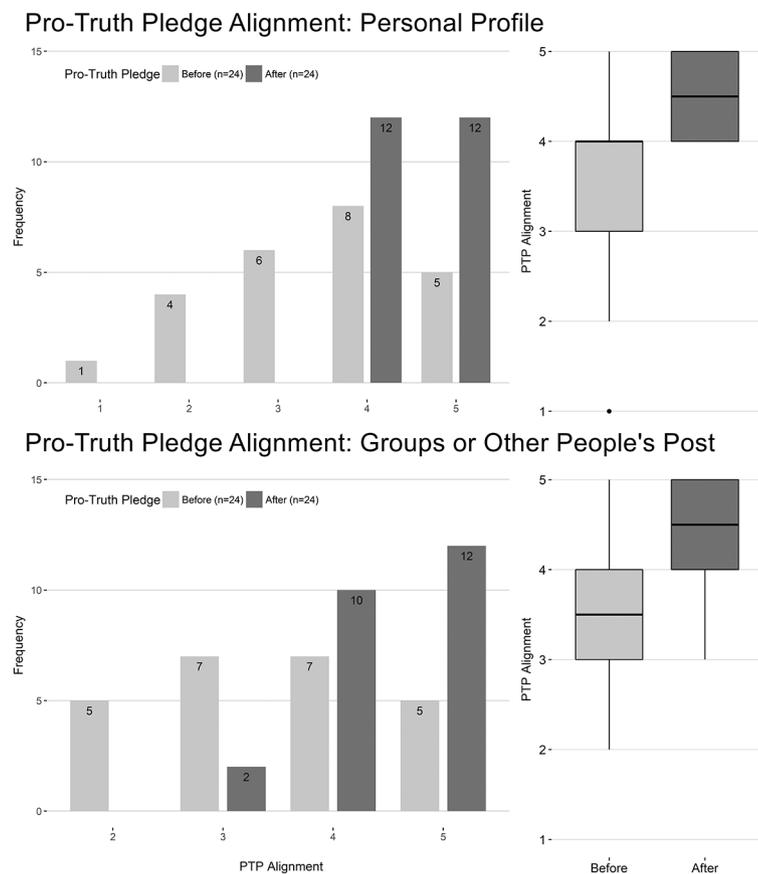


Figure 1. Visual summary of preliminary survey data with PTP alignment in Facebook engagement.

Much more research is needed on whether the findings of this preliminary survey data actually corresponds to people's real behavior, and also replication is needed with a larger and randomly assigned sample with a control group. Likewise, the preliminary survey data was gathered from Americans, and we would like to conduct a study with a more geographically diverse population in the future. While we have no reason to suspect that the PTP would be more or less likely to impact Americans as opposed to citizens of other countries, it would be valuable to investigate this question. The only meaningful distinction in the impact of the pledge on Americans might come from the fact that there is a stronger incentive to target Americans by foreign actors disseminating fake news, combined with a lack of meaningful action by the US government in addressing the problem of fake news. We are intending to pursue these studies in the future.

## Conclusion

To solve the problem of fake news sharing, we need technocognitive solutions, meaning ones that combine technology with psychological principles, according to prominent researchers in the field. The Pro-Truth Pledge, which combines psychology research with online mechanisms of implementation and propagation, and crowdsources fact-checking, is one such potential intervention. It asks participants to commit to twelve behaviors, which are intended to counteract a number of cognitive biases that contribute to people believing in and sharing misin-

formation, an essential aspect of the psychology research informing the content of the pledge itself. The full pledge reads as follows:

I Pledge My Earnest Efforts To:

**Share** the truth

- Verify: fact-check information to confirm it is true before accepting and sharing it
- Balance: share the whole truth, even if some aspects do not support my opinion
- Cite: share my sources so that others can verify my information
- Clarify: distinguish between my opinion and the facts

**Honor** the truth

- Acknowledge: acknowledge when others share true information, even when we disagree otherwise
- Reevaluate: reevaluate if my information is challenged, retract it if I cannot verify it
- Defend: defend others when they come under attack for sharing true information, even when we disagree otherwise
- Align: align my opinions and my actions with true information

**Encourage** the truth

- Fix: ask people to retract information that reliable sources have disproved even if they are my allies
- Educate: compassionately inform those around me to stop using unreliable sources even if these sources support my opinion
- Defer: recognize the opinions of experts as more likely to be accurate when the facts are disputed
- Celebrate: celebrate those who retract incorrect statements and update their beliefs toward the truth

In addition to committing to the behaviors, pledge-takers are encouraged to share about taking the pledge on their social media, to put markers of taking the pledge on their online profiles, and to fact-check other pledge-takers, which is the crowd-sourcing component of the pledge.

This technocognitive solution has shown some early signs of effectiveness. To be effective requires that: 1) we see evidence of people taking the pledge, and 2) we see evidence of people abiding by the behaviors of the pledge. We have clear evidence of people taking the pledge when they find out about it through venues like coverage by prominent media, through personal outreach, through word-of-mouth interactions on social media, and other means. The evidence on people abiding by the behaviors of the pledge is much weaker, and only suggestive rather than conclusive. Initial qualitative evidence includes case studies of interviews and observations, while quantitative evidence includes a preliminary survey with self-selection and self-reporting by participants. The pledge organizers plan to undertake future quantitative studies to evaluate whether people do indeed abide by the pledge after taking it. In the meantime, the early data suggests the pledge is valuable, and while more research on its impact is needed, it may be beneficial to encourage the widespread adoption of this techno-cognitive tool for the sake of fighting the insidious problem of fake news.

## Funding

The authors have no funding to report.

## Competing Interests

The authors have declared that no competing interests exist.

## Acknowledgments

The Ohio State University provided facilities and Institutional Review Board oversight.

## References

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *The Journal of Economic Perspectives*, *31*, 211-236. doi:10.1257/jep.31.2.211
- The Annenberg Public Policy Center of the University of Pennsylvania. (2017). Jamieson offers new name for fake news: 'Viral Deception' or VD. Retrieved from <http://www.annenbergpublicpolicycenter.org/on-cnn-jamieson-offers-new-name-for-fake-news-viral-deception-or-v-d>
- Ariely, D., & Jones, S. (2012). *The (honest) truth about dishonesty: How we lie to everyone—Especially ourselves*. New York, NY, USA: HarperCollins.
- Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, *13*, 219-224. doi:10.1111/1467-9280.00441
- Barrera, O., Guriev, S., Henry, E., & Zhuravskaya, E. (2017). *Facts, alternative facts, and fact checking in times of post-truth politics* (CEPR Discussion Paper No. 12220). Retrieved from <https://ssrn.com/abstract=3018474>
- Booth, R., Weaver, M., Hern, A., & Walker, S. (2017, November 14). "Russia used hundreds of fake accounts to tweet about Brexit, data shows. *The Guardian*. Retrieved from <https://www.theguardian.com/world/2017/nov/14/how-400-russia-run-fake-accounts-posted-bogus-brexit-tweets>
- Brückner, H., & Bearman, P. (2005). After the promise: The STD consequences of adolescent virginity pledges. *The Journal of Adolescent Health*, *36*, 271-278. doi:10.1016/j.jadohealth.2005.01.005
- Burn, S. M. (1991). Social psychology and the stimulation of recycling behaviors: The block leader approach. *Journal of Applied Social Psychology*, *21*, 611-629. doi:10.1111/j.1559-1816.1991.tb00539.x
- Cilizza, C. (2016, November 2). How the heck can voters think Donald Trump is more honest than Hillary Clinton? *Washington Post*. Retrieved from [https://www.washingtonpost.com/news/the-fix/wp/2016/11/02/donald-trump-hasnt-told-the-truth-repeatedly-in-this-campaign-voters-still-think-he-is-more-honest-than-hillary-clinton/?utm\\_term=.0f490d1b4201](https://www.washingtonpost.com/news/the-fix/wp/2016/11/02/donald-trump-hasnt-told-the-truth-repeatedly-in-this-campaign-voters-still-think-he-is-more-honest-than-hillary-clinton/?utm_term=.0f490d1b4201)
- The Collins Word of the Year 2017 is.... (2017). In *Collins English Dictionary*. Retrieved from <https://www.collinsdictionary.com/woty>
- Connolly, K., Chrisafis, A., McPherson, P., Kirchgaessner, S., Haas, B., Phillips, D., . . . Safi, M. (2016, December 2). Fake news: An insidious trend that's fast becoming a global problem. *The Guardian*. Retrieved from <https://www.theguardian.com/media/2016/dec/02/fake-news-facebook-us-election-around-the-world>

- Correia, V., & Festinger, L. (2014). *Biased argumentation and critical thinking*. Bern, Switzerland: Peter Lang.
- Dietz, T., Ostrom, E., & Stern, P. C. (2003). The struggle to govern the commons. *Science*, *302*, 1907-1912. doi:10.1126/science.1091015
- Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. *Advances in Experimental Social Psychology*, *44*, 247-296. doi:10.1016/B978-0-12-385522-0.00005-6
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, *105*, 98-121. doi:10.1016/j.obhdp.2007.05.002
- Enfield, N. (November 16, 2017). We're in a post-truth world with eroding trust and accountability. It can't end well. *The Guardian*. Retrieved from <https://www.theguardian.com/commentisfree/2017/nov/17/were-in-a-post-truth-world-with-eroding-trust-and-accountability-it-cant-end-well>
- European Commission. (2017, November 13). *Next steps against fake news: Commission sets up High-Level Expert Group and launches public consultation* [Press Release]. Retrieved from [http://europa.eu/rapid/press-release\\_IP-17-4481\\_en.htm](http://europa.eu/rapid/press-release_IP-17-4481_en.htm)
- Fang-yuan, L. U. (2007). Evolutionary game analysis on environmental pollution problem. *Systems Engineering-Theory & Practice*, *9*, 148-152.
- Farand, C. (2017, April 22). French social media awash with fake news stories from sources 'exposed to Russian influence' ahead of presidential election. *The Independent*. Retrieved from <http://www.independent.co.uk/news/world/europe/french-voters-deluge-fake-news-stories-facebook-twitter-russian-influence-days-before-election-a7696506.html>
- Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, *144*, 993-1002. doi:10.1037/xge0000098
- Frijda, N. H., Manstead, A. S., & Bem, S. (Eds.). (2000). *Emotions and beliefs: How feelings influence thoughts*. New York, NY, USA: Cambridge University Press.
- Garrett, R. K., & Weeks, B. E. (2013, February). The promise and peril of real-time corrections to political misperceptions. In *Proceedings of the 2013 Conference on Computer supported Cooperative Work* (pp. 1047-1058). doi:10.1145/2441776.2441895
- Gino, F., Norton, M. I., & Ariely, D. (2010). The counterfeit self: The deceptive costs of faking it. *Psychological Science*, *21*, 712-720. doi:10.1177/0956797610366545
- GoFundMe.com. (2017). Fundraiser by Michael Smith: Michael W. Smith for congress 2018. Retrieved from <https://www.gofundme.com/MichaelWSmith2018>
- Gottfried, J., & Shearer, E. (2016, May 26). News use across social media platforms 2016. *Pew Research Center*. Retrieved from <http://www.journalism.org/2016/05/26/news-use-acrosssocial-media-platforms-2016>
- Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: Status, reputation, and conspicuous conservation. *Journal of Personality and Social Psychology*, *98*, 392-404. doi:10.1037/a0017346
- Green, J., & Issenberg, S. (2016, October 27). Inside the Trump bunker, with days to go. *Bloomberg Businessweek*. Retrieved from <https://www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go>

- Guadagno, R. E., & Cialdini, R. B. (2010). Preference for consistency and social influence: A review of current research findings. *Social Influence, 5*, 152-163. doi:10.1080/15534510903332378
- Guillory, J. J., & Geraci, L. (2013). Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition, 2*, 201-209. doi:10.1016/j.jarmac.2013.10.001
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY, USA: Pantheon Books.
- Hanley, N., & Folmer, H. (1998). *Game theory and the environment*. Cheltenham, United Kingdom: Edward Elgar Publishing.
- Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243-1248. doi:10.1126/science.162.3859.1243
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science, 2*, 96-100. doi:10.1111/1467-8721.ep10770953
- Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology, 69*, 1069-1086. doi:10.1037/0022-3514.69.6.1069
- Hopper, J. R., & Nielsen, J. M. (1991). Recycling as altruistic behavior: Normative and behavioral strategies to expand participation in a community recycling program. *Environment and Behavior, 23*, 195-220. doi:10.1177/0013916591232004
- Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., . . . Weber, U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters, 23*, 487-504. doi:10.1007/s11002-012-9186-1
- Jolls, C., Sunstein, C. R., & Thaler, R. (1998). A behavioral approach to law and economics. *Stanford Law Review, 50*, 1471-1550. doi:10.2307/1229304
- Katzev, R. D., & Pardini, A. U. (1987). The comparative effectiveness of reward and commitment approaches in motivating community recycling. *Journal of Environmental Systems, 17*, 93-113. doi:10.2190/XV00-DD4B-EPEH-EN5R
- Kray, L. J., & Galinsky, A. D. (2003). The debiasing effect of counterfactual mind-sets: Increasing the search for disconfirmatory information in group decisions. *Organizational Behavior and Human Decision Processes, 91*, 69-81. doi:10.1016/S0749-5978(02)00534-4
- Kruger, J., & Dunning, D. (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*, 1121-1134. doi:10.1037/0022-3514.77.6.1121
- Kwong, J. (2017, November 18). "Why Clinton lost: What Russia did to control the American mind and put Trump in the White House. *Newsweek*. Retrieved from <http://www.newsweek.com/how-russian-trolls-targeted-2016-election-merry-go-round-lies-715912>
- Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "Post-Truth" Era. *Journal of Applied Research in Memory and Cognition, 6*, 353-369. doi:10.1016/j.jarmac.2017.07.008
- Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science, 4*, 390-398. doi:10.1111/j.1745-6924.2009.01144.x
- Lord, K. R. (1994). Motivating recycling behavior: A quasi experimental investigation of message and source strategies. *Psychology and Marketing, 11*, 341-358. doi:10.1002/mar.4220110404
- Mann, H., Garcia-Rada, X., Houser, D., & Ariely, D. (2014). Everybody else is doing it: Exploring social transmission of lying behavior. *PLoS One, 9*, Article e109591. doi:10.1371/journal.pone.0109591

- Martino, S. C., Elliott, M. N., Collins, R. L., Kanouse, D. E., & Berry, S. H. (2008). Virginity pledges among the willing: Delays in first intercourse and consistency of condom use. *The Journal of Adolescent Health, 43*, 341-348. doi:10.1016/j.jadohealth.2008.02.018
- Mazar, N., Amir, O., & Ariely, D. (2008a). More ways to cheat – Expanding the scope of dishonesty. *JMR, Journal of Marketing Research, 45*, 651-653.
- Mazar, N., Amir, O., & Ariely, D. (2008b). The dishonesty of honest people: A theory of self-concept maintenance. *JMR: Journal of Marketing Research, 45*, 633-644. doi:10.1509/jmkr.45.6.633
- Mazar, N., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Journal of Public Policy & Marketing, 25*, 117-126. doi:10.1509/jppm.25.1.117
- McCabe, D. L., & Trevino, L. K. (1993). Academic dishonesty: Honor codes and other contextual influences. *The Journal of Higher Education, 64*, 522-538.
- McCabe, D. L., Trevino, L. K., & Butterfield, K. D. (1999). Academic integrity in honor code and non-honor code environments: A qualitative investigation. *The Journal of Higher Education, 70*, 211-234.
- McDermott, R. (2004). The feeling of rationality: The meaning of neuroscientific advances for political science. *Perspectives on Politics, 2*, 691-706. doi:10.1017/S1537592704040459
- Milinski, M., Semmann, D., & Krambeck, H. J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature, 415*, 424-426. doi:10.1038/415424a
- Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*, 103-122. doi:10.1002/ejsp.2420220202
- Myers, T. A., Maibach, E. W., Roser-Renouf, C., Akerlof, K., & Leiserowitz, A. A. (2013). The relationship between personal experience and belief in the reality of global warming. *Nature Climate Change, 3*, 343-347. doi:10.1038/nclimate1754
- Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A., & Nielsen, R. K. (2017). *Reuters Institute digital news report 2017*. Retrieved from SSRN repository: <https://ssrn.com/abstract=3026082>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175-220. doi:10.1037/1089-2680.2.2.175
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior, 32*, 303-330. doi:10.1007/s11109-010-9112-2
- Nyhan, B., & Reifler, J. (2015). The effect of fact-checking on elites: A field experiment on US State Legislators. *American Journal of Political Science, 59*, 628-640. doi:10.1111/ajps.12162
- Oskamp, S., Harrington, M. J., Edwards, T. C., Sherwood, D. L., Okuda, S. M., & Swanson, D. C. (1991). Factors influencing household recycling behavior. *Environment and Behavior, 23*, 494-519. doi:10.1177/0013916591234005
- Ostrom, E. (2015). *Governing the commons*. New York, NY, USA: Cambridge University Press.
- Palmer, E. (2017, November 18). Spain Catalonia: Did Russian 'fake news' stir things up? *BBC News*. Retrieved from <http://www.bbc.com/news/world-europe-41981539>
- Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2013). Belief bias during reasoning among religious believers and skeptics. *Psychonomic Bulletin & Review, 20*, 806-811. doi:10.3758/s13423-013-0394-3

- Pew Research Center. (2016, May 26). News use across social media platforms 2016. *Pew Research Center*. Retrieved from <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016>
- Peer Reviewers' Openness Initiative. (2017). Frequently Asked Questions. Retrieved from <https://opennessinitiative.org/frequently-asked-questions>
- Pro-Truth Pledge. (2017a). Pro-Truth Pledge Search Engine. Retrieved from <https://www.protruthpledge.org/facts-search-engine>
- Pro-Truth Pledge. (2017b). Threading the fact-checking needle. Retrieved from <https://www.protruthpledge.org/threading-fact-checking-needle>
- Rasmussen Reports. (2016, September 30). Voters don't trust media fact-checking. *Rasmussen Reports*. Retrieved from [http://www.rasmussenreports.com/public\\_content/politics/general\\_politics/september\\_2016/voters\\_don\\_t\\_trust\\_media\\_fact\\_checking](http://www.rasmussenreports.com/public_content/politics/general_politics/september_2016/voters_don_t_trust_media_fact_checking)
- Schwarzer, R. (2008). Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. *Applied Psychology, 57*, 1-29.
- Selinger, E., & Whyte, K. (2011). Is there a right way to nudge? The practice and ethics of choice architecture. *Sociology Compass, 5*, 923-935. doi:10.1111/j.1751-9020.2011.00413.x
- Sheldon, O. J., Dunning, D., & Ames, D. R. (2014). Emotionally unskilled, unaware, and uninterested in learning more: Reactions to feedback about deficits in emotional intelligence. *The Journal of Applied Psychology, 99*, 125-137. doi:10.1037/a0034138
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *SIGKDD Explorations, 19*, 22-36. doi:10.1145/3137597.3137600
- Shu, L. L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 15197-15200. doi:10.1073/pnas.1209746109
- Shuster, S. (2017, August 9). Russia has launched a fake news war on Europe: Now Germany is fighting back. *Time*. Retrieved from <http://time.com/4889471/germany-election-russia-fake-news-angela-merkel>
- Silverman, C. (2016, November 16). This analysis shows how fake election news stories outperformed real news on Facebook. *BuzzFeed News*. Retrieved from [https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm\\_term=.lsyJR21epR#.edLdEVoz5E](https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm_term=.lsyJR21epR#.edLdEVoz5E)
- Silverman, C., & Singer-Vine, J. (2016, December 6). Most Americans who see fake news believe it, new survey says. *BuzzFeed News*. Retrieved from [https://www.buzzfeed.com/craigsilverman/fake-news-survey?utm\\_term=.bxym5zP245#.argg0PIBD0](https://www.buzzfeed.com/craigsilverman/fake-news-survey?utm_term=.bxym5zP245#.argg0PIBD0)
- Smith, M. [Michael]. (2017, May 24). Due to a Truth Pledge I have taken I have to say I have not been able to verify this post [Imgur post]. Retrieved May 24, 2017, from <http://imgur.com/a/A8IOY>
- Smola, J. (2017, October 23). Ohio State professor's 'Pro-Truth Pledge' encourages fact-checking before sharing information. *The Columbus Dispatch*. Retrieved from <http://www.dispatch.com/news/20171023/ohio-state-professors-pro-truth-pledge-encourages-fact-checking-before-sharing-information>
- Subramanian, S. (2017, February 15). Inside the Macedonian fake-news complex. *Wired*. Retrieved from <https://www.wired.com/2017/02/veles-macedonia-fake-news>
- Sunstein, C. R., & Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review, 70*, 1159-1202. doi:10.2307/1600573

- Swire, B., Berinsky, A. J., Lewandowsky, S., & Ecker, U. K. (2017). Processing political misinformation: Comprehending the Trump phenomenon. *Royal Society Open Science*, 4, Article 160802. doi:10.1098/rsos.160802
- Swift, A. (September 14, 2016). Americans' trust in mass media sinks to new low. *Gallup*. Retrieved from <http://www.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx>
- Thaler, R. H., & Sunstein, C. R. (1999). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT, USA: Yale University Press.
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *The American Economic Review*, 93, 175-179. doi:10.1257/000282803321947001
- Van Lange, P. A., De Bruin, E. M., Otten, W., & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology*, 73, 733-746. doi:10.1037/0022-3514.73.4.733
- Van Vugt, M. (2009). Averting the tragedy of the commons: Using social psychological science to protect the environment. *Current Directions in Psychological Science*, 18, 169-173. doi:10.1111/j.1467-8721.2009.01630.x
- Van Vugt, M. (2001). Community identification moderating the impact of financial incentives in a natural social dilemma: Water conservation. *Personality and Social Psychology Bulletin*, 27, 1440-1449. doi:10.1177/01461672012711005
- Van Vugt, M., & Samuelson, C. D. (1999). The impact of personal metering in the management of a natural resource crisis: A social dilemma analysis. *Personality and Social Psychology Bulletin*, 25, 735-750. doi:10.1177/0146167299025006008
- Verkuyten, M., & Nekuee, S. (1999). Ingroup bias: The effect of self-stereotyping, identification and group threat. *European Journal of Social Psychology*, 29, 411-418. doi:10.1002/(SICI)1099-0992(199903/05)29:2/3<411::AID-EJSP952>3.0.CO;2-8
- Vining, J., & Ebreo, A. (1992). Predicting recycling behavior from global and specific environmental attitudes and changes in recycling opportunities. *Journal of Applied Social Psychology*, 22, 1580-1607. doi:10.1111/j.1559-1816.1992.tb01758.x
- Vraga, E. K., Tully, M., & Rojas, H. (2009). Media literacy training reduces perception of bias. *Newspaper Research Journal*, 30, 68-81. doi:10.1177/073953290903000406
- WEF. (2013). Outlook on the global agenda 2014. *WEF World Economic Forum*. Retrieved from <http://reports.weforum.org/outlook-14>
- Westman, M., Eden, D., & Shirom, A. (1985). Job stress, cigarette smoking and cessation: The conditioning effects of peer support. *Social Science & Medicine*, 20, 637-644. doi:10.1016/0277-9536(85)90402-2
- Word of the year 2016 is.... (2016). In *Oxford Dictionaries*. Retrieved from <https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016>
- Zimmerman, R. S., & Connor, C. (1989). Health promotion in context: The effects of significant others on health behavior change. *Health Education Quarterly*, 16, 57-75. doi:10.1177/109019818901600108