Solving Fake News and Other Misinformation: The Pro-Truth Pledge

Gleb Tsipursky
Board Chair, Intentional Insights, Columbus OH
gleb@intentionalinsights.org

We have witnessed an alarming deterioration of truth in the US public sphere, especially in the political arena. This paper describes a proposed intervention, the Pro-Truth Pledge, which combines behavioral science with crowd-sourcing to help address this problem. The pledge asks signees – private citizens and public figures – to commit to 12 behaviors that research in behavioral science shows correlate with an orientation toward truthfulness. Pledge mechanisms like this one have been shown in other contexts to lead private citizens to engage in more pro-social behavior. For public figures, the pledge offers specific incentives to stick to the pledge, with rewards in the form of positive reputation for honesty and truth-telling, and accountability through evaluation and potential punishment for deception.

keywords: truth, fake news, honest, politics, citizens, politicians

1. Introduction

Few would dispute that many people have lied to achieve their political agendas in the past, but this problem has become particularly bad lately. Recent political events, such as the successful tactics used by Donald Trump's campaign 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."[1]

On the one hand, post-truth political methods have to do with the quantity of lies. For example, *The Washington Post*'s well-respect Fact-Checking Column has compared the two major candidates in the US presidential election in early November 2016, and found

that one of them – Trump – received their worst rating for claims fact-checking 63 percent of the time, while the other candidate – Hillary Clinton – received the worst rating 14.2 percent of the time (previously, most candidates received the worst rating between 10 and 20 percent of the time).[2]

On the other, post-truth politics involves a new model of behavior when caught lying. Unlike previous politicians who backed away when caught lying, post-truth politicians do not back away from their falsehoods. Instead, they attack those who point out their deceptions, undermining public trust in credible experts and reliable news sources. 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.[3]

This is not only a problem with public figures: fake news, more recently termed "viral deception" by Kathleen Hall Jamieson, director of the Annenberg Public Policy Center, is sweeping social media, shared by ordinary citizens.[4] 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.[5]

We propose an intervention we term the Pro-Truth Pledge (PTP). The purpose of the PTP is to change the incentive structure for public figures, establishing rewards for truth-oriented activities and penalties for sharing misinformation. It targets politicians and other public figures, with the intent of committing them to spreading accurate information as informed by behavioral science research. Another goal is to get ordinary citizens to take the pledge, and to advocate for others and especially public figures to sign the pledge, as well as to monitor them for violations. The pledge draws on behavioral science research that takes advantage of findings on rewards and punishments, reputation management, and choice architecture.

2. 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.[6] 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.[7] 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.[8] 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.[9] The environmental movement presents many examples of successful efforts to addressing the tragedy of the commons in

environmental pollution.[10] Only substantial disincentives for polluting outweigh the benefits of polluting from a game theoretical perspective.[11] Particularly illuminating is a theoretical piece by Mark van Vugt describing the application of behavioral sciences insights 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.[12]

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 behavioral science about problematic human thinking patterns – cognitive biases – to shape human behavior for social good while also respecting individual freedom of choice.[13] 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.[14]

3. A Proposed Intervention to Address Pollution of Truth: The Pro-Truth Pledge

The Pro-Truth Pledge (PTP), created by a team of behavioral scientists, 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 behavioral science research informing the content of the pledge itself.[15]

> I Pledge 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.[16] 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.[17] 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.

Misinformation is anything that goes against the truth of reality. It can mean directly lying about the situation at hand, for instance when an athlete denies taking steroids that she was actually taking. It can mean lying by omission, as when a scholar publishes a study with a successful experiment, while hiding that he conducted 50 of the same experiments that failed, until by random chance one finally worked, a phenomenon known as publication bias. In some cases, misinformation is obvious, so that anyone can see it. In other cases, it is less so. For those cases, the PTP calls on pledge signers to rely on credible fact-checking websites and/or on the scientific consensus.

Rather than going through the process of vetting fact-checking websites, we have decided to outsource that work to Facebook, which is partnering with websites it has vetted and evaluated as credible. As of the initial unveiling, the websites include Snopes, Politifact, ABC News, and FactCheck.org, and more will be added over time. All these are members of a common coalition, the Poynter International Fact Checking Network, and have committed to a common set of principles. Any other websites that Facebook uses will be considered credible for PTP purposes. Someone who takes the pledge will be considered in violation of the pledge if they make a claim that is similar to those rated as "mostly false" or "completely false" by one of these websites (they use different language, but you get the idea). In a case where credible websites disagree, for instance one calls a claim "mostly false" and another calls it "mostly true," we will not consider the claim a violation of the PTP.

In some cases, fact-checking websites have not evaluated certain claims, but the claim will be opposed by scientific research. Since science is the best of all methods we as human beings have found to determine the reality about the world and predict the outcomes of our actions, someone will be evaluated as in violation of the pledge if they make a claim that goes against the scientific consensus. We are comfortable with the Wikipedia definition of scientific consensus as "the collective judgment, position, and opinion of the community

[4]

of scientists in a particular field of study. Consensus implies general agreement, though not necessarily unanimity. Consensus is normally achieved through communication at conferences, the publication process, replication (reproducible results by others), and peer review. These lead to a situation in which those within the discipline can often recognize such a consensus where it exists, but communicating to outsiders that consensus has been reached can be difficult, because the 'normal' debates through which science progresses may seem to outsiders as contestation. On occasion, scientific institutes issue position statements intended to communicate a summary of the science from the 'inside' to the 'outside' of the scientific community." Thus, we can recognize scientific consensus by position statements by prestigious scientific organizations or the result of meta-analysis studies (evaluations of a series of other prominent studies) that come to a clear determination.

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.[18]

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.[19]

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.[20] 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, in other words are ignorant of their own ignorance.[21] 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.[22] 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.[23] 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. Thus, any proposed solution needs to address the perception of dishonesty by others and oneself, and also address benefits to one's in-group from dishonesty.

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.[24] 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.[25] Our likelihood of lying is strongly impacted by our social network, making it especially important to address social norms around deception.[26] Dan Ariely summarizes and synthesizes the research on what moves us to lie and vice versa in his *The Honest Truth About Dishonesty: How We Lie to Everyone---Especially Ourselves*. In a nutshell, he finds 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.[27] The more of these factors an solution can address, the better.

4. The Pro-Truth Pledge: Private Citizens

We separate the targets for the pledge signees into two categories, private citizens and public figures, and will talk about the former first. Why would private citizens take the pledge? Many people 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.[28] Furthermore, signing the pledge gives any individual who signs it greater credibility among their peers who know they signed it. The pledge encourages individuals who signed it to share about it on their social media and personal networks, and also put a badge on their online presence indicating they signed it.[29] They get access to unique resources available to

signees, such as a search engine composed of credible sources verified as reliable by the PTP organizers.[30] 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.[31] 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 behavioral science research on the environmental movement showed that people who report self-identification with a community tend to engage in behaviors condoned by that community.[32]

However, would pledge-takers who are private citizens, and thus have no external monitoring, follow such behaviors upon taking the pledge? Behavioral science research on precommitment suggests that those who commit to a certain behavioral norm will be more likely to follow it.[33] 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.[34] A related phenomena is a preference for consistency, which recent research suggests influences many people to make decisions that are consistent with their past decisions.[35] Most relevantly for the PTP, at schools that have honor codes students tend to engage in less academic dishonesty.[36] Likewise, signing an honor code before a test tends to decrease cheating compared to signing an honor code at the end of a test.[37] 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.[38] 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.[39] Anecdotal evidence from current pledge signees (over 1900 so far) shows some self-reported impact.[40]

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

[7]

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.[41] 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.[42] 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.[43] 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.[44] 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.[45] Active volunteering and community engagement in recycling programs, such as block-leader programs, proved even more effective in increasing recycling behavior.[46] 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.

5. The Pro-Truth Pledge: Public Figures

Why should public figures take the PTP? We anticipate that some public figures would be motivated by the same intrinsic motivations that would lead private citizens to take the pledge. However, we wanted to provide particular incentives for public figures to take the pledge, and also disincentives for breaking the pledge, and we decided to do so in the form of reputation. Reputational rewards and penalties have been shown to be vital in addressing tragedies of the commons in the environmental movement.[47] Other research also demonstrated the social benefits of coordinated punishments to sustain cooperation and prevent defection.[48] The PTP borrows from this approach.

How are public figures rewarded for taking the pledge? Taking the pledge is a way of providing credible information about the honesty of a public figure to an audience interested in such information, thus providing a substantial reputational reward. When signing the pledge, each public figure has an opportunity to provide a brief statement about why they took the pledge, and some links to their online presence. This information will be stored in a publicly-accessible database that anyone can access, such as constituents

interested in evaluating political candidates for office or deciding whether to trust the commentary of a media figure, policy expert, or academic commenting on public affairs. Moreover, the statement would get sent in a regular newsletter to all pledge signers who chose to subscribe to email updates. Doing so improves that public figure's reputation and gains them new supporters. The public figure can provide additional content for the PTP newsletter about how the pledge changed their behavior, further reinforcing both their reputation and providing proof for the PTP newsletter subscribers of the effectiveness of the pledge, creating a virtuous cycle characteristic of successful innovations.[49]

Such provision of information has been crucial in successful interventions within the environmental movement to address tragedies of the commons. As an example, research shows 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.[50] Similarly, many consumers of political information lack knowledge about which officials and media figures and analysts are credible, and the PTP pledge provides that information.

Many may worry about the problem of false signaling or cheating – a public figure may take the pledge to signal a commitment to the truth, without actually abiding by the pledge.[51] Private citizens have little incentive to take their time and share their personal data by filling out the pledge, making it likely that only those committed to advancing the cause of truth in our society would take this action. However, the reputational value for public figures of taking the pledge, especially as the PTP gains popularity and credibility and also has a bigger email list, will grow higher and higher. If we do not prevent false signaling and cheating on the pledge, the pledge will not be able to provide credible information and thus fail to shift incentives to favor sharing accurate information instead of deception.

To address cheating, the pledge involves a monitoring mechanism that makes sure the pledge has teeth in the form of reputational penalties which are commensurate with the infraction. Some PTP advocates are assigned the duty of monitoring public figures. If an advocate suspects that a public figure violated the pledge, the advocate will contact the individual privately, with an approach of "innocent until reasonably shown guilty" perspective – perhaps the person misspoke, or the advocate misunderstood. If the public figure withdraws the statement, or the advocate finds no likely violation of the pledge, the matter ends there.

If the advocate still thinks there might be a violation of the pledge, the advocate will then escalate the matter to PTP mediating committee, depending on the stature of the public figure. While anyone who signs up to the PTP may become an advocate, mediating committees are composed of a group of vetted volunteers who will evaluate the evidence provided by the advocate, contact the public figure for a chance to offer a defense, and make a ruling. If there is a ruling of a violation, then this ruling is evaluated by a member of the PTP Central Coordinating Committee, to ensure fairness and accuracy, and provide an external perspective. In the case that the PTP Central Coordinating Committee member also determines that a violation has occurred, the committee will then contact the public figure, offering the person a final chance to retract the statement. If the public figure still refuses to take their words back, the PTP mediating committee will then consider that the public figure

has made a deliberate decision to lie, and will rule the public figure to be in contempt of the pledge.

This process might sound a little convoluted, but it minimizes the possibility of the PTP being politicized or corrupted at a local level, a concern raised by many in the formulation of the pledge. Indeed, research on the environmental movement showed that for an institution such as the PTP to succeed in gaining trust and credibility, it needs to demonstrate transparent, clear, and fair rules and procedures where all participants have a chance to make their case and feel heard. For instance, research on the California water shortage in 1991 showed that people cooperated with drastic water-saving measures by local water authorities only if they believed these authorities to listen to the concerns of all and provide clear, accurate, and unbiased information.[52]

Once someone is found to be in contempt of the pledge, the mediating committee will then proceed to put reputational pressure on the individual to get that individual to change their position on the matter. It would issue a press advisory to all relevant media – for instance, all the media in the San Francisco area if the public figure is the mayor of San Francisco – that the public figure is in contempt of the pledge. It will also issue an action alert to those who indicated they want to receive such alerts – either at the local, regional, or national level, depending on the stature of the public figure – for them to email, tweet, call, write, and protest in front of the office of the public figure encouraging the person to revise the relevant statement, and writing letters-to-the-editor about the situation. Finally, the public figure will be listed on the PTP website as in contempt of the pledge.

We anticipate that these consequences will provide considerable reputation pressure for a public figure to avoid being in contempt of the pledge. If the public figure envisions violating the pledge deliberately, they would be better off not signing it at all. Thus, the pledge is not simply cheap talk, as it has strong reputational pressure behind it. We already have an example of how a violation of the pledge was resolved. Michael Smith, a candidate for Congress from Idaho, took the Pro-Truth Pledge.[53] 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."[54] He indicated that he would be more careful with future postings.

So why should elected or appointed officials take the pledge if it restrains their activities and causes them to make such statements retracting their posts? Officials need to be perceived as trustworthy by citizens. The PTP provides that credibility, due to the presence of the monitoring mechanism. Citizens can easily look them up in the PTP database. If the official has signed the pledge a while ago and is not in contempt, the citizen can assume the official has not made any deceptive statements without retracting them later. The official gets additional benefits because when the official signs up, her information is included in the PTP updates. This provides the official with positive reputation as being honest and credible, and gets them more support. There is an additional benefit for elected officials whose opponent for office has not taken the PTP, since the official can raise

questions about why the opponent does not wish to take the pledge. The PTP thus offers a first mover advantage for those public officials who take it early onward.[55] Politicians are already taking the pledge, with over 650 having done so already, showing its promise as a tool to shift incentives.[56]

What about policy experts, commentators, analysts, media figures, and scholars? They all need to be perceived as trustworthy by the audiences to which they communicate. The PTP provides them with that benefit due to the monitoring mechanism, and similarly to the officials described above, the longer they are signed up without being in contempt, the more credibility they get. Those who sign can also get a broader audience engaged with them since their information will be included in the PTP updates. Moreover, if their competitors do not sign the pledge, those who signed up will get a bigger audience, since audiences will start flocking to those deemed more trustworthy sources of news/analysis/thought leadership. Thus, the first mover advantage applies to these groups as well. Media figures are also taking the pledge, for example a conservative radio and podcast host, John Wells.[57]

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.[58] This distrust for fact-checkers will not be solved by providing more fact-checking, and 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.

Of course, the Pro-Truth Pledge may not work despite the problems with the current best alternatives. Regarding private citizens, virginity pledges have been shown consistently to delay the onset of sexual behavior.[59] 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.[60] 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

[11]

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.

6. Conclusion

The PTP uses all four components shown by behavioral science research on environmental pollution as crucial to addressing tragedies of the commons.[61] 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 and reputational penalties for dishonesty, taking advantage of the behavioral science research on incentives. By learning from the successes of the environmental movement, the Pro-Truth Pledge can help address the pollution of truth in our public sphere, especially in politics.

You can held advance this cause by going to ProTruthPledge.org, taking the pledge, encouraging those in your social networks to do so, and calling on your elected representatives to pledge to truthfulness. Although it is very unlikely that readers of this article need a pledge to exhibit the truthful behaviors of the pledge, it is very likely that at least some people in your social network do, and it's certainly the case for at least some of your elected representatives. That's why globally-known public intellectuals such as Peter Singer, Steven Pinker, Jonathan Haidt, and others took the pledge, and so did secular humanist activists like Herb Silverman, Aron Ra, Dale MacGowan, Dan Barker, Ed Brayton, Noah Lugeons, and many more. Join them at ProTruthPledge.org to solve the crisis of misinformation destroying democracies around the world!

Notes

1. Oxford University Press. (n.d.). "Word of the Year 2016 is ... post-truth." http://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016
2. Cilizza, C. (2016, November 2). "How the heck can voters think Donald Trump is more honest than Hillary Clinton?" *Washington Post*, http://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
3. Swift, A. (2016, September 14). "Americans' Trust in Mass Media Sinks to New Low." http://www.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx

4. The Annenberg Public Policy Center of the University of Pennsylvania. (2017, March 6). "Jamieson Offers New Name for Fake News: 'Viral Deception' or VD." http://www.annenbergpublicpolicycenter.org/on-cnn-jamieson-offers-new-name-for-fake-news-viral-deception-or-v-d/

5. Nyhan, B., and Reifler, J. (2010). "When corrections fail: The persistence of political misperceptions." *Political Behavior* 32(2): 303-330. McDermott, R. (2004). "The feeling of rationality: The meaning of neuroscientific advances for political science." *Perspectives on Politics* 2(4): 691-706. Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion.* New York: Pantheon Books.

6. Hardin, G. (1968). "The tragedy of the commons." *Science* 162(3859): 1243-1248.

7. Change, R. (2000). *The Global Commons: Environmental and Technological Governance*, 2nd edn. New York: Wiley.

8. Hanley, N., and Folmer, H. (1998). *Game Theory and the Environment*. Cheltenham, UK: Edward Elgar Publishing.

9. Hardin, G. (1968). "The tragedy of the commons." *Science* 162(3859): 1243-1248.

10. Ostrom, E. (2015). *Governing the Commons*. Cambridge, UK: Cambridge University Press. Feeny, D., Berkes, F., McCay, B. J., and Acheson, J. M. (1990). "The tragedy of the commons: twenty-two years later." *Human Ecology* 18(1): 1-19. Berkes, Fikret. "Fishermen and 'the tragedy of the commons'." *Environmental Conservation* 12 (1985): 199-206.

11. Fang-yuan, L. U. (2007). "Evolutionary game analysis on environmental pollution problem [J]." *Systems Engineering–Theory and Practice* 9: 148-152.

12. Van Vugt, Mark. (2009). "Averting the tragedy of the commons: Using social psychological science to protect the environment." *Current Directions in Psychological Science* 18(3): 169-173.

13. Thaler, R. H., and Sunstein, C. R. (2003) "Libertarian paternalism is not an oxymoron." *The University of Chicago Law Review* 70(4): 1159-1202; Thaler, R. H., and Sunstein, C. R. (2003). "Libertarian paternalism." *The American Economic Review* 93(2): 175-179. These are further elaborated in specific prescriptions and approaches to public policy in a popular trade book that might be of particular relevance to policy experts: Thaler Richard, H., and Sunstein Cass, R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, Conn.: Yale University Press.

14. Thaler, R. H., Sunstein, C. R., and Balz, J. P. (2013). "Choice architecture." In *The Behavioral Foundations of Public Policy*, ed. Eldar Shafer, 428-439. Princeton, N.J.: Princeton University Press, 2013. Jolls, C., Sunstein, C. R., and Thaler, R. (1998). "A behavioral approach to law and economics." *Stanford Law Review* 50(5): 1471-1550. Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., and Wansink, B. (2012). "Beyond nudges: Tools of a choice architecture." *Marketing Letters* 23(2): 487-504. Selinger, E., and Whyte, K. (2011). "Is there a right way to nudge? The practice and ethics of choice architecture." *Sociology Compass* 5(10): 923-935.

15. Pro-Truth Pledge. (n.d.). Pro-Truth Pledge. http://www.protruthpledge.org/

16. Nickerson, R. S. (1998). "Confirmation bias: A ubiquitous phenomenon in many guises." *Review of General Psychology* 2(2): 175.

17. Kray, L. J., and Galinsky, A. D. (2003). "The debiasing effect of counterfactual mind-sets: Increasing the search for disconfirmatory information in group decisions." *Organiza-tional Behavior and Human Decision Processes* 91(1): 69-81. Hirt, E. R., and Markman, K. D. (1995). "Multiple explanation: A consider-an-alternative strategy for debiasing judgments." *Journal of Personality and Social Psychology* 69(6): 1069; Lilienfeld, S. O., Ammirati, R., and Landfield, K. (2009). "Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare?" *Perspectives on Psychological Science* 4(4): 390-398.

18. Pro-Truth Pledge (n.d.). Pro-Truth Pledge FAQs. Retrieved from https://www.protruthpledge.org/

19. Allcott, H., and Gentzkow, M. (2017). *Social Media and Fake News in the 2016 Election* (No. w23089). National Bureau of Economic Research.

20. Mullen, B., Brown, R., and Smith, C. (1992). "Ingroup bias as a function of salience, relevance, and status: An integration." *European Journal of Social Psychology* 22(2): 103-122. Verkuyten, M., and Nekuee, S. (1999). "Ingroup bias: the effect of self-stereotyping, identification and group threat." *European Journal of Social Psychology* 29(2-3): 411-418.

21. Dunning, D. (2011). "The Dunning-Kruger Effect: On Being Ignorant of One's Own Ignorance." *Advances in Experimental Social Psychology* 44: 247. Kruger, J., and 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(6): 1121. Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., and Kruger, J. (2008). "Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent." *Organizational Behavior and Human Decision Processes* 105(1): 98-121. Sheldon, O. J., Dunning, D., and Ames, D. R. (2014). "Emotionally unskilled, unaware, and uninterested in learning more: Reactions to feedback about deficits in emotional intelligence." *Journal of Applied Psychology* 99(1): 125.

22. Gino, F., Norton, M. I., and Ariely, D. (2010). "The counterfeit self: The deceptive costs of faking it." *Psychological Science* 21(5): 712-720.

23. Mazar, N., and Ariely, D. (2006). "Dishonesty in everyday life and its policy implications." *Journal of Public Policy and Marketing* 25(1): 117-126.

24. Mazar, N., Amir, O., and Ariely, D. (2008). "The dishonesty of honest people: A theory of self-concept maintenance." *Journal of Marketing Research* 45(6): 633-644. Mazar, N., Amir, O., and Ariely, D. (2008). "More ways to cheat – expanding the scope of dishonesty." *Journal of Marketing Research* 45(6): 651-653.

25. Katzev, R. D., and Pardini, A. U. (1987). "The comparative effectiveness of reward and commitment approaches in motivating community recycling." *Journal of Environmental Systems* 17(2): 93-113.

26. Mann, H., Garcia-Rada, X., Houser, D., and Ariely, D. (2014). "Everybody else is doing it: Exploring social transmission of lying behavior." *PloS One* 9(10): e109591.

27. Ariely, D., and Jones, S. (2012). *The (Honest) Truth about Dishonesty: How We Lie to Everyone, Especially Ourselves*. New York: HarperCollins.

28. Van Lange, P.A.M., De Bruin, E., Otten, W., and Joireman, J.A. (1997). "Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence." *Journal of Personality and Social Psychology* 73: 733-746. Van Vugt, M., and Samuelson, C.D. (1999). "The impact of metering in a natural resource crisis: A social dilemma analysis." *Personality and Social Psychology Bulletin* 25: 731-745.

29. Pro-Truth Pledge (n.d.). Pro-Truth Pledge badge. Retrieved from https://www.protruthpledge.org/website-badge-seal/

30. Pro-Truth Pledge (n.d.). Pro-Truth Pledge search engine. Retrieved from https://www.protruthpledge.org/facts-search-engine/

31. Zimmerman, R. S., and Connor, C. (1989). "Health promotion in context: the effects of significant others on health behavior change." *Health Education and Behavior* 16(1): 57-75. Westman, M., Eden, D., and Shirom, A. (1985). "Job stress, cigarette smoking and cessation: the conditioning effects of peer support." *Social Science and Medicine* 20(6): 637-644.

32. Van Vugt, M. (2001). "Community identification moderating the impact of financial incentives in a natural social dilemma: A water shortage." *Personality and Social Psychology Bulletin* 27: 1440-1449.

33. Ariely, D., and Wertenbroch, K. (2002). "Procrastination, deadlines, and performance: Self-control by precommitment." *Psychological Science* 13(3): 219-224.

34. Correia, V., and Festinger, L. (2014). "Biased argumentation and critical thinking." In *Rhetoric and Cognition: Theoretical Perspectives and Persuasive Strategies.* Bern, Switz.: Peter Lang.

35. While many people do have a preference for consistency, others do not have such a preference, so this factor will influence only a portion of pledge-takers: Guadagno, R. E., and Cialdini, R. B. (2010). "Preference for consistency and social influence: A review of current research findings." *Social Influence* 5(3): 152-163.

36. According to studies of self-reported cheating by students. See McCabe, D. L., and Trevino, L. K. (1993). "Academic dishonesty: Honor codes and other contextual influences." *The Journal of Higher Education* 64(5): 522-538. McCabe, D. L., Trevino, L. K., and Butterfield, K. D. (1999). "Academic integrity in honor code and non-honor code environments: A qualitative investigation." *Journal of Higher Education* 70(2): 211-234.

37. Shu, L. L., Mazar, N., Gino, F., Ariely, D., and 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* 109(38): 15197-15200.

38. The intervention was for a period of eight weeks: Katzev, R. D., and Pardini, A. U. (1987). "The comparative effectiveness of reward and commitment approaches in motivating community recycling." *Journal of Environmental Systems* 17(2): 93-114.

39. Frijda, N. H., Manstead, A. S., and Bem, S. (2000). *Emotions and Beliefs: How Feelings Influence Thoughts*. Cambridge, UK: Cambridge University Press. Ariely, D., and

[15]

Jones, S. (2012). *The (Honest) Truth about Dishonesty: How We Lie to Everyone, Especially Ourselves*. New York: HarperCollins.

40. Here is a statement from one pledge signee on his motivations for taking the pledge and the impact of the pledge on his behavior:

> I'm a retired US intelligence agent, having worked diligently for 4 decades to ensure that our country stayed safe and secure. I won't tell you my name (my former employers would probably frown on that), but you can call me Bill. Born in the post-WWII American heartland to a blue-collar, Protestant family tracing its American roots back to the earliest Colonial days, I was also a volunteer in the U.S. Army during Vietnam. As with my intelligence career, I joined out of a patriotic duty to contribute to the security of my country. And as a patriot I still hold to our basic American values, among them frankness, honesty in all my dealings, a belief that government is here to serve the people and not the reverse, and a deep commitment to telling the truth. Thus, unsurprisingly, I was immediately taken by the ideas and ideals behind the Pro-Truth Pledge. In this age of rampant duplicity and sell-outs in politics, the media, Hollywood, professional sports, even our religious institutions; the idea of a group of people dedicated to finding, spreading and sticking to the truth—even when it contradicts their most closely-held beliefs— felt like being offered a glass of ice-cold lemonade after a day of working in the summer sun. Since we're being honest here, I've got as many cherished beliefs and 'stone tablets' as the next guy. But with the PTP I was being offered an opportunity to walk-the-talk. So I signed the pledge.

> It took no time at all before my resolve to follow the pledge was tested. A political piece (no need to go into the details here) that played right to my particular political biases hit cable TV and then the Internet and of course my first inclination was to share it as quickly and widely as possible. But then I remembered the pledge I'd signed and put the brakes on. I decided to wait a bit to see how it played out (and boy-howdy am I glad I did.) As things unfolded I started seeing things that didn't add up. So I began using various fact-checking sites to see what they might have to add to the story. As it turned out the story was a complete dud, "fake news" as they say. That experience has led me to be much more vigilant in assessing, and sharing, stories that appeal to my political sensibilities. I now make a much bigger effort to fact-check before I post or share.

> Fortunately, the folks responsible for the PTP have made this much easier by putting together a Google-based custom search engine that combines a number of fact-checking algorithms into one. Called the "Pro-Truth Pledge Fact Search", you can find it at: http://www.protruthpledge.org/facts-search-engine/ Give it a test-drive and see what you think. Intentional Insights, the organization running the pledge, and the PTP, inspired me to vow not to respond in anger to any web posts that challenge what I may say or share. Rather, I strive to listen, consider and respond gently and with humility, in a way that disarms confrontations and opens all parties up to dialogue.

[16]

41. Schwarzer, R. (2008). "Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors." *Applied Psychology* 57(1): 1-29.

42. Oskamp, S., Harrington, M. J., et al. (1991). "Factors influencing household recycling behavior." *Environment and Behavior* 23(4): 494-519.

43. Vining, J., and Ebreo, A. (1992). "Predicting recycling behavior from global and specific environmental attitudes and changes in recycling opportunities." *Journal of Applied Social Psychology* 22(20): 1580-1607.

44. Lord, K. R. (1994). "Motivating recycling behavior: A quasi-experimental investigation of message and source strategies." *Psychology and Marketing* 11(4): 341.

45. Griskevicius, V., Tybur, J. M., and Van den Bergh, B. (2010). "Going green to be seen: status, reputation, and conspicuous conservation." *Journal of Personality and Social Psychology* 98(3): 392.

46. Hopper, J. R., and Nielsen, J. M. (1991). "Recycling as altruistic behavior normative and behavioral strategies to expand participation in a community recycling program." *Environment and Behavior* 23(2): 195-220. Burn, S. M. (1991). Social psychology and the stimulation of recycling behaviors: The block leader approach." *Journal of Applied Social Psychology* 21(8): 611-629.

47. Milinski, M., Semmann, D., and Krambeck, H. J. (2002). "Reputation helps solve the 'tragedy of the commons'." *Nature* 415(6870): 424-426.

48. Boyd, R., Gintis, H., and Bowles, S. (2010). "Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare." *Science* 328(5978): 617-620.

49. Casadesus-Masanell, R., and Ricart, J. E. (2011). "How to design a winning business model." *Harvard Business Review* 89(1-2): 100-107.

50. Dietz, T., Ostrom, E., and Stern, P.C. (2003). "The struggle to govern the commons." *Science* 302: 1907-1912.

51. Connelly, B. L., Certo, S. T., Ireland, R. D., and Reutzel, C. R. (2011). "Signaling theory: A review and assessment." *Journal of Management* 37(1): 39-67.

52. Tyler, T.R., and Degoey, P. (1995). "Collective restraint in social dilemmas: Procedural justice and social identification effects on support for authorities." *Journal of Personality and Social Psychology* 69: 482-497.

53. GoFundMe (n.d). GoFundMe fundraiser for Michael W Smith for Congress 2018. http://www.gofundme.com/MichaelWSmith2018

54. Imgur (2017, May 24). Imgur photograph of Michael Smith's public post. Retrieved from http://imgur.com/a/A8IOY

55. Kerin, R. A., Varadarajan, P. R., and Peterson, R. A. (1992). "First-mover advantage: A synthesis, conceptual framework, and research propositions." *The Journal of Marketing* 56(4): 33-52.

56. For instance, here is a statement from one such politician, the Democrat Dan Epstein:

> As a progressive who has always valued learning to make our society better, as a Democrat who believes in ethics and transparency in government and

politics, as a lifelong student and teacher who has always been devoted to the sciences, humanities, and all forms of study, I will tell the truth, promote the truth, and live the truth. I will stand against not only my opponents, but my own co-partisans if need be, to honor the truth in the face of falsehood. I am running for the US House of Representatives in the Texas 19th Congressional District in 2018. http://www.danepsteinforwesttexas.com/

Here is another one, from Republican Jay Baumeister:

> I feel it is time to bring the country back together and this can not be done the way congress is acting now in an us vs them mentality. Most congressmen have only one goal and that is to get reelected. Congressmen will say whatever they need to in order to accomplish that goal truth or not. I pledge to work toward the truth and to be willing to speak the truth even if it is not in my best interest politically. I am a Republican running for Congress in Ohio's District 12: http://www.facebook.com/Baumeister-for-Congress-1682557778660008/

These statements are perfect illustrations of how politicians from both sides of the aisle aim to get reputational benefits from taking the pledge. The statements are publicly available on the PTP website and will be sent to all who took the pledge and signed up for updates: https://www.protruthpledge.org/public-figures-signed-pledge/

57. Here is his statement for the PTP newsletter on taking the pledge:

> The lifeblood of my program to which my name is attached and therefore all who I call and who call me, friend, those who trust me to be honest with them, and most importantly in the Earthly realm, my family rely on truthfulness in what I do. And of supreme importance, God is watching. And listening. www.caravantomidnight.com

Similarly to the politicians, Wells' statement is designed to get him appropriate reputational rewards. Since Wells announced his commitment to take the pledge on his program, his listeners are now holding him accountable, along with PTP advocates who are assigned to this task.

58. Rasmussen Reports (2016, September 30) "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

59. Martino, S. C., Elliott, M. N., Collins, R. L., Kanouse, D. E., and Berry, S. H. (2008). "Virginity pledges among the willing: Delays in first intercourse and consistency of condom use." *Journal of Adolescent Health* 43(4): 341-348.

60. Brückner, H., and Bearman, P. (2005). "After the promise: the STD consequences of adolescent virginity pledges." *Journal of Adolescent Health* 36(4): 271-278.

61. Van Vugt, Mark. (2009). "Averting the tragedy of the commons: Using social psychological science to protect the environment." *Current Directions in Psychological Science* 18(3): 169-173; Milinski, M., Semmann, D., and Krambeck, H. J. (2002). "Reputation helps solve the 'tragedy of the commons'." *Nature* 415(6870): 424-426.