The detrimental effect of low political truth literacy on democratic systems

Jack Harich¹, Montserrat Koloffon Rosas²
¹Thwink.og, Atlanta, USA. jack@thwink.org. Corresponding author.
²Vrije Universiteit Amsterdam, Amsterdam, Netherlands.

Abstract

After a steady transition to more and stronger democracies, a precipitous backward slide to authoritarianism/autocracy has begun. The paper explores why this has occurred and how the backward slide can be reversed. A system dynamics model was developed to forge a path beyond existing theories of “why democracies break down.” Results indicate the main root cause is low political truth literacy. As long as this is low democracy cannot function as intended because citizens are too easily deceived into voting against their own best interests. A sample solution element for raising political truth literacy, Truth Literacy Training, was empirically tested. Study results indicate average political truth literacy is currently low and can be raised to high with a surprisingly small amount of carefully designed training, though a collection of solution elements is likely required for optimal root cause resolution.

Introduction

Instead of the permanent ascendancy of Western liberal democracy as the final form of government famously predicted by Fukuyama (1992) in The End of History and the Last Man, democratic backsliding toward authoritarianism/autocracy is strongly underway (Bermeo, 2016; Lueders and Lust, 2018; Waldner and Lust, 2018). Many scholars have concluded that liberal democracy is in crisis (Benkler et al., 2018, p. 4; Wiesner et al., 2019).

The European Union, once a showcase of the benefits of democracy, finds itself in an “existential crisis” and has “reached a point where liberal democracy relapses throughout the continent” (Onis and Kutlay, 2019). After an early turn toward democracy after the collapse of the USSR in 1991, Russia has veered toward dictatorship under Vladimir Putin and is now rated as an autocratic regime (V-Dem-Institute, 2021). China, a one-party state which formerly elected new leaders periodically, in 2018 become a dictatorship with the constitutional removal of president term limits, allowing Xi Jinping to remain in power for life. The United States, after four years under Donald Trump’s embrace of authoritarian ideals and attack of democratic institutions using a new style of authoritarian populism (Chacko and Jayasuriya, 2016), only narrowly re-
turned to democratic ideals under Joe Biden, by 44,000 out of a total of 74 million votes in the 2020 election due to use of an electoral college system (Swasey and Hanzhang, 2020). However, American’s steep slide toward authoritarianism continues (IDEA, 2021; Kagan, 2021). India, Turkey, and Israel, “all previously hailed as exceptional democracies,” have all turned to antidemocratic populism (Rogenjofer and Panievsky, 2020). The problem has grown so acute that the number one agenda item for the December 2021 Summit for Democracy, attended by 89 nations, is “Defending against authoritarianism” (SummitForDem, 2021).

Using their Liberal Democracy Index, the latest V-Dem Democracy Report (V-Dem-Institute, 2021), found: “The global decline during the past 10 years is steep and continues in 2020, especially in the Asia-Pacific region, Central Asia, Eastern Europe, and Latin America. The level of democracy enjoyed by the average global citizen in 2020 is down to levels last found around 1990.”

This decline has critical implications for common good problems like climate change because “The connections between the widespread rise of authoritarian and populist leaders …and destructive trends in environmental politics and governance on the other are legion” (McCarthy, 2019). The rise of authoritarian regimes also increases geopolitical tensions and the likelihood of armed conflict (Brands, 2018).

However, “we lack theories to explain backsliding, though we have long engaged in a perhaps interminable debate about the causes of democratic transitions, democratic breakdowns, authoritarian resilience, and democratic consolidation” (Waldner and Lust, 2018). Specifically, “there is a lack of work on the question of under which circumstances might instability arise” (Wiesner et al., 2019).

The topic of “why democracies break down” has drawn “huge amounts of attention” from scholars (Bermeo, 2016). Waldner and Lust (2018) provide an overview of the literature by examination of six theory families developed to explain why democratic backsliding occurs. The theories center on political agency, political culture, political institutions, political economy, social structure and political coalitions, and international actors. Each offers a different set of loosely organized factors that could logically cause backsliding. None offer a rigorous cohesive theory. Waldner and Lust conclude that “despite the existence of six well-populated theory families, we do not have an obvious theoretical framework for explaining backsliding.” Fukuyama (2020) himself acknowledges a “democratic recession” is underway and must be reversed, but offers no advice allowing practical reversal. Instead, his research is limited to recommending further investigation into the relationship between democracy, corruption, and state capacity.

The literature thus offers no theoretical explanation of democratic backsliding. The paper addresses this gap by reporting on two research questions: (1) What is the key dynamic structure behind observed declines in liberal democracy and its main root causes? (2) Can it be empirically demonstrated, on a preliminary basis, that the root causes exist and can be resolved?

The first question was addressed by system dynamics modeling. The main root cause of the decline was found to be low political truth literacy. Political truth literacy is
the ability to tell truth from falsehood in political statements meant to influence voters and to make correct voting decisions based on that information.

The second question was addressed by designing the Truth Literacy Training solution element and an empirical study. Results confirm the main root cause exists (political truth literacy is currently low) and that a person’s political truth literacy can be raised from low to high with a short amount of training. A follow up study found the training effect persists, with a small decline. This decline is easily corrected with a short amount of refresh training.

The paper presents an analysis explaining the decline and how it can be reversed by pushing on the right high leverage points. A sample solution element for doing this, Truth Literacy Training, is then described and tested. The paper ends with discussion.

**Analysis of the democratic backsliding problem**

**Summary of analysis results**

In a chapter titled *The Principle of Leverage*, Peter Senge (1990) states that: “The bottom line of systems thinking is leverage—seeing where actions and changes in structures can lead to significant, enduring improvements (p114). ... As a systems thinker, you would first identify the key problem symptoms, and then the symptomatic and fundamental responses to it (p120).” Forrester (1971, pp. 94–95, italics added) describes the same two responses:

People are often led to intervene at points in a system where little leverage exists…. ...social systems are inherently insensitive to most policy changes that people select in an effort to alter behavior. In fact, a social system draws attention to the very points at which an attempt to intervene will fail. Human experience, which has been developed from contact with simple systems, leads us to look close to the symptoms of trouble for a cause. But when we look, we are misled because the social system presents us with an apparent cause that is plausible according to the lessons we have learned from simple systems, although this apparent cause is usually a coincident occurrence that, like the trouble symptom itself, is being produced by the feedback loop dynamics of a larger system.

Accordingly, Figure 1 summarizes analysis results by emphasizing the symptomatic and fundamental responses, the key feedback loop, the causal chain from symptoms to root cause, and the leverage points. The diagram is organized into two layers: (1) the superficial (symptomatic) layer of the problem, where intermediate causes are so easy to see they are routinely assumed to be root causes, and (2) the deeper fundamental layer, where by understanding the problem’s structure its root causes may be found. Forrester’s “apparent cause” is what the diagram calls an intermediate cause. “Little leverage exists” if people assume the apparent cause is the root cause because that leads to pushing on low leverage points, where “an attempt to intervene will fail.” Only by pene-
trating to the fundamental layer with tools like system dynamics can “the feedback loop dynamics of a larger system” be discovered.

Figure 1. High-level causal diagram of the analysis using standard systems thinking and root cause terms, and the three high-level forces at play: S, F, and R. A simpler but less accurate term for the main root cause is low political truth literacy.

That pushing on a low leverage point cannot resolve an intermediate cause is described on the diagram with S<R, meaning superficial solution forces (S) are always less than root causes forces (R). Only fundamental solutions that resolve root causes can solve a causal problem. This is described with F>R, meaning fundamental solutions can work because fundamental solution forces (F) can be greater than root cause forces (R). The three high-level forces are what must be correctly understood to effectively solve the problem and to provide a comprehensive theory explaining backsliding.

The first why question

Root causes are found by applying some form of the Five Whys method (Imai, 1986, p. 50), where starting at problem symptoms, the analyst asks “WHY does this occur?” until the root causes are found. This reveals the causal chain (and the feedback loops structure in more complex problems) running from symptoms to intermediate causes to one or more root causes. For difficult problems this requires asking why many times. A root cause is the deepest cause in a causal chain (or the most basic cause in a
feedback loop structure) that can be resolved. Resolved means the problem will probably not recur due to that root cause.

Our first question was WHY is backsliding from democracy to authoritarianism occurring? Svolik (2019) reports that 197 democratic backslides occurred from 1973 to 2018. Of these, 46 were military coups and 88 were executive takeovers via election, with takeovers averaging about 80% of all backslides after the end of the Cold War in 1991. In an examination of How Democracies Die, Levitsky and Ziblatt (2018, p. 5) summarize this change: “Democratic backsliding today begins at the ballot box.”

Thus, the main first intermediate cause of backsliding is election of politicians not working for the democratic common good, represented by percent degenerates infected in Figure 2. Too many citizens are voting instead for politicians working for the uncommon good of powerful special interests. These may be a single person (a dictator or authoritarian) or a ruling elite (a party, aristocracy, theocracy, oligarchy, ruling family, etc.). If backsliding is underway, today the special interest is usually an authoritarian.

The solution used now is the same solution used before to hasten the spread of democracy before the backslide began. The solution attempts to promote and prove the superiority of democracy over authoritarianism. This is a form of more of the truth, a low leverage point, represented by the true memes node in Figure 2.

As modeled by Harich (2010), “more of the truth” is the traditional general-purpose process used to solve public interest problems. Labeled Classic Activism, the process has four steps: 1. Identify the problem, 2. Find the truth about what should be done to solve the problem, 3. Promote the truth, and if that doesn’t work, 4. Magnify the truth via exhortation, inspiration, and bargaining. On difficult problems the process frequently fails, usually because change resistance is high or because since the process lacks consideration of root causes, problem solvers remain confined to the superficial layer of the problem, as is the case here.

The more people accept the truth of the superiority of democracy over all other forms of government, the more they tend to elect politicians who seek to strengthen democracy. For example, Halperin et. al. (2010, p. xvii and xi) reviewed fifty years of data for all developing countries (including China) and found “that democracies have compelling advantages over their authoritarian counterparts in fostering social and economic development.” The data “provides a cogent, well-documented refutation of the dictator-is-best nostrum.” Texts like Halperin’s are used by writers, international banks, development agencies, governments, etc. to promote the truth about the superiority of democracy.

The chief reason the solution no longer works is that since about 2000, authoritarian state capacity (government effectiveness, regulatory quality, rule of law, and control of corruption) has improved so much worldwide, especially in China, that there is no longer proof democracy is superior (Foa, 2018, p. 133). “Of the twenty fastest growing countries of the past two decades, fifteen have been autocratic regimes. Of the fifteen wealthiest economies in the world today by per capita income, almost two-thirds are nondemocracies.” Halperin’s data is not current enough to show this change. Fukuya-
ma’s prediction in 1992 that liberal democracy had permanently proven itself to be superior to all other forms of government has, at least in the short term, failed.

The second why question

Our next question was WHY does election of politicians not working for the common good occur? The answer is mainly because of successful political deception, represented by undetected false memes in Figure 2. Other factors like voter suppression, voter fraud, rigged elections, bribery, etc. are minor. Deception is the main technique used to convince an electorate majority to act against their own best interests.

The evidence shows that all authoritarians seeking to gain office, as well as those in office, depend heavily on deceptively provoking a wide range of false beliefs and emotions, especially fear. McCarthy (2019) describes how:

They use bellicose rhetoric and gestures in theatrical efforts to project strength. They promise to take quick and decisive action on highlighted issues, in contrast to liberal democratic administrations portrayed as weak, passive, and indecisive. They make the central populist move of claiming to speak and act in the name of and with the support of “the people,” who are typically identified in nativist, xenophobic, and often explicitly racialized terms. Following closely from that, they often identify internal enemies—ethnic or religious minorities, immigrants, refugees, drug users—as scapegoats and targets for public anger. They use populist rhetorical tropes of resentful anti-elitism, suspicion of experts and complexity, and celebration of direct action to promise simple, immediate solutions to complex, long-term problems.

Once in power, all authoritarian regimes rely on deceptive strategies and copious amounts of state-managed propaganda to maintain public support. Walker (2016) explains why:

Above all, authoritarian rulers are preoccupied with regime survival, and they study and learn from other authoritarian regimes, both past and present, in order to maintain power. … Through experimentation and learning, authoritarian regimes have refined their techniques of manipulation at the domestic level. By constructing fake political parties, phony social movements, and state-controlled media enterprises that appear in many ways to be like those of their democratic counterparts, autocrats simulate democratic institutions as a way of preventing authentic democracy from taking root. …authoritarians can deploy a potent combination of censorship and propaganda, allowing them to dominate the media space and create an unchallenged alternate reality for their audiences.

Solutions used to solve the problem of successful political deception employ a strategy of misinformation correction, a low leverage point also represented by true memes in Figure 2. If people believe statements that are not true, then that can be corrected by providing citizens with corrected versions of deceptive statements pointing out the truth. Solutions to do this are fact-checks, articles, social media posts, news, etc. pointing out the truth.
Unfortunately, these superficial solutions have worked so poorly that we are now living in the post-truth age of politics, where: “The post-truth politician does not simply pick-and-choose among relevant facts, offer questionable interpretations or avoid inconvenient questions. The post-truth politician manufactures his or her own facts. The post-truth politician asserts whatever they believe to be in their own interest and they continue to press those same claims, regardless of the evidence amassed against them” (Lockie, 2017). Notable recent examples are the “policy deception” behind the disastrous Brexit vote (Baines et al., 2020), Donald Trump’s deception-based election in 2016, his 30,573 false or misleading claims while in office (Kessler et al., 2021), and “The Big Lie” promoted by Trump and supporters that he lost the 2020 election only because of massive election fraud. Despite lack of evidence, 68% of US Republicans, 26% of independents, and 6% of Democrats believe the election was stolen from Trump (PPRI_Staff, 2021).

The third why question and the simulation model

Digging deeper, WHY does successful political deception occur? This was so difficult to answer we constructed a system dynamics model (Figure 2). The reason for the difficulty was we had completed our work on the easy-to-see superficial layer of the problem and were now attempting to understand the fundamental layer, where for difficult problems causal structure is so hard to see correctly that further analysis often requires additional tools, such as simulation modeling. Diagraming the superficial layer (Figure 1), we could see the intermediate causes were part of a reinforcing loop. But we could not see deeper detail without constructing a model.
The model uses Dawkins’ (1976) concept of memes. A *meme* is copied information capable of affecting behavior. All memes are learned from others, either directly from other people or indirectly through a transmission medium such as books, television, or social media. In the model a meme is a statement that is true or false. When a meme enters a person’s mind, they are said to be “infected” with the meme.

The backbone of the model is the three stocks and two dueling loops. The upper stock contains supporters infected by false memes. The middle stock contains uninfected supporters. The lower stock contains supporters infected by true memes. The dueling loops determine movement between stocks.

The Dueling Loops model follows the long tradition of small system dynamics models designed to communicate powerful counterintuitive insights to the public and policy makers (Ghaffarzadegan et al., 2011), such as URBAN1 (38 nodes). Because the
Dueling Loops model (37 nodes) is easily understood and exhibits clear behavior, “important insights regarding the source of policy failures can be uncovered.”

Small insight models entail estimated parameters for archetypical/exploratory use or measured parameters for actual cases. Using estimation, the Dueling Loops model was tuned to give realistic behavior over the full range of the high leverage points, false meme size, and influence per degenerate or rationalist.

The model and later study use two important acronyms. LTQ (*logical truth quotient*) is the ability to logically tell if a deceptive claim is true or false. DTQ (*democratic truth quotient*, aka political truth literacy) is the ability to vote correctly given a deceptive statement made by a politician. LTQ is an input to DTQ, so theoretically DTQ can never exceed LTQ.

The Race to the Bottom among Politicians (the key feedback loop driving the problem) works in this manner: *Supporters Due to Degeneration* use their *degenerates influence* to transmit *false memes*. LTQ and *false meme size* determines how many *false memes* become *detected false memes*. DTQ times *detected false memes* gives *actionable false memes*. The theory that DTQ cannot exceed LTQ is enforced by use of arrow A on the model. *False memes* minus *actionable false memes* gives *undetected false memes*. These cause the *degenerates infectivity rate*, which after a delay causes the *degenerates maturation rate*. This causes some *Not Infected Neutralists* to move to the stock of *Supporters Due to Degeneration*, and the loop starts over again.

Opposed to the Race to the Bottom is the Race to the Top. Here, instead of attracting supporters with falsehoods in an attempt to support the desires of powerful special interests, politicians attract them with the objective truth about what’s best for the common good. The two loops are mirror opposites with one critical difference. The Race to the Top lacks an *undetected true memes* node, because there is no deception to detect. The two main loops are locked in a perpetual struggle to attract supporters.

This structure models the essence of the left-right political spectrum, consequential because “global politics is first and foremost a debate between the left and the right. ... The left-right dichotomy occupies a special place, as the most enduring, universal, and encompassing of all political strategies” (Noel and Therien, 2008, p. 3).

The Dueling Loops capture this dichotomy. The two loops each embody an enduring political strategy. The Race to the Top houses the liberal/progressive left, who believe in democracy, equality, justice, freedom of many kinds, and quality of life because that optimizes the common good. The Race to the Bottom houses the conservative/authoritarian right, who promote anything that maximizes what powerful special interests want. In reality there are many loops along the political spectrum, such as the far right, right, center, etc. We have modeled only two.

Examples of powerful special interests are authoritarians, managers of large for-profit corporations, the rich, and elite ruling groups of many kinds, e.g., the ruling class. All are a small percentage of the electorate. In a democracy, the only way a minority can persuade the majority to vote for special interest candidates is by use of deception, favoritism, threats/intimidation, or lethal force. Force (Such as don’t show up on election day or you will be shot.) is illegal. Threats/intimidation (Like vote for my candidate
or suffer the consequences!) are mostly illegal and largely ineffective since balloting is secret. Favoritism, such as bribery or patronage, is mostly illegal. Furthermore, in a large population, favoritism can only attract a small percentage of supporters. It’s too expensive to bribe millions of voters. There are not enough perks and jobs to dole out to millions of people.

This explains why deception is the main strategy for Race to the Bottom politicians and why political deception in democratic systems is so ubiquitous. Jeremy Bentham, the father of utilitarianism, in his handbook of political fallacies published in 1824 reached the same conclusion: “…it is impossible by fair reasoning ...to justify the sacrifice of the interests of the many to the interests of the few.... It follows that for effecting this purpose they must have recourse to every kind of fallacy, and address themselves, when occasion requires it, to the passions, the prejudices, and the ignorance of mankind” (Larrabee, 1925, p. xxi). Referring to Western democracies, “The central problem facing conservatives, once their country’s [voting] franchise had been extended to include most adult men, was that it was unclear why most voters would want to vote for them” (Ware, 1996, p. 32).

The key model insight is that the size of a lie, and hence its supporter attractive power, can be inflated, while the size of the truth cannot. The bigger the lie, the bigger and stronger the false belief a deceptive politician can plant in a supporter’s mind. The Race to the Bottom contains an inherent advantage that the Race to the Top lacks.

For example, a virtuous politician may gain supporters by stating, “I know we can’t balance the budget any time soon, but I will form a panel of experts to determine what the best we can do is.” Meanwhile, a deceptive politician is garnering supporters by saying, “Economics is easy. You just put a firm hand on the tiller and go where you want to go. I can balance the budget in four years, despite what the experts are saying. They are just pundits. Don’t listen to them. A vote for me is a vote for a better future.” The deceptive politician is also telling numerous different groups, “Yes, I can do that for you. No problem.” Guess who will usually win?

From a mathematical perspective, the size (and hence the appeal) of a falsehood can be inflated by saying that 2 + 2 = 5, or 7, or even 27, but the size of the truth can never be inflated by saying anything more than 2 + 2 = 4. Inflation is used to create fear when there is nothing to fear, doubt when there is nothing to doubt, the false promise of I can do so-and-so for you when I really cannot, a large flaw in one’s opponent when there is only a small flaw or no flaw, and so on.

We can now answer our third why question: WHY does successful political deception occur? The answer is because of the inherent advantage of the Race to the Bottom, represented by false meme size in Figure 2. Since it appears this advantage can be resolved in a practical manner by pushing on the high leverage point of raise political truth literacy from low to high in Figure 1, it is the main root cause. For simplification we usually say the root cause is low political truth literacy.

The single high leverage point in the high-level causal diagram becomes two detailed high leverage points in the simulation model: (1) The first is LTQ (logical truth quotient). Raising it increases detected false memes, which increases actionable false
memes. (2) The second high leverage point is DTQ (democratic truth quotient). Raising it increases actionable false memes, which weakens the Race to the Bottom by decreasing undetected false memes and strengthens the Race to the Top by increasing desertion.

Dynamic behavior of the model

This section demonstrates how the two high leverage points behave with eleven simulation runs (Figure 3). Percent rationalists = Supporters Due to Rationality / (Supporters Due to Rationality + Supporters Due to Degeneration). Runs 1 to 5 begin with 1 degenerate supporter, 1 rationalist, and 98 neutralists, as well as zero initial percent degenerates and rationalists infected. Runs 6 to 11 begin with the model in equilibrium and last 100 instead of 500 years. In all runs the two influence variables are equal and never changed.

LTQ and DTQ are the high leverage points. LTQ is the ability to spot political deception, which causes detected false memes. DTQ is the ability to spot political deception and then use that information to vote correctly, which causes actionable false memes. If DTQ is zero, then actionable false memes is zero, regardless of LTQ.

Run 1. This shows how when neither side (rationalists and degenerates) has an advantage, percent rationalists stays unchanged at 50%. The number of rationalists and degenerates rise evenly. Neither side has an advantage since false meme size = 1 and logical truth quotient (LTQ) = 0. Demographic truth quotient (DTQ) doesn’t matter because detected false memes = 0.

Run 2. False meme size is raised from 1 to 1.1. While telling tiny lies offers only a small advantage, over time it accumulates into a large one. After a long period of time, degenerates vastly outnumber rationalists.

Run 3. Let’s give the other side a small advantage. False meme size is moved back to 1, LTQ is raised from zero to 20%, and DTQ remains zero. Nothing happens, because LTQ only increases detected false memes. Because DTQ is zero, so is actionable false memes. The results of run 3 are identical to run 1. In the next run let’s increase DTQ also.

Run 4. DTQ is raised from 0 to 20%. Now it’s the rationalists who have the advantage. They quickly become the strongly dominant political group, with percent rationalists at the end of the simulation run rising to 97%.

Because DTQ is required to increase actionable false memes and DTQ can never theoretically exceed LTQ, LTQ and DTQ are equal in runs 4 and up. Any LTQ greater than DTQ has no effect.

Run 5. Reference mode 500 years. Social agents are adaptive. Degenerate politicians are clever enough to adjust the size of lies to the optimum size: not too big and not too small. The effect of size of lie on detection is a lookup table with a curve to reflect how once a lie becomes too big, it’s more easily detected and diminishing returns begin. Experimentation shows the optimum false meme size is 2.9. Now the degenerates are the ones with the advantage. Percent rationalists flattens out at 39%. We hypothesize that in most political systems both high leverage points are low, at about 20%. This run
thus reflects approximate real-world behavior and is the reference mode. It is the problem to solve, as later runs attempt to do.

Figure 3. Simulation runs 1 to 11. NA is not applicable.
Because political truth literacy is low, the Race to the Bottom is the dominant loop most of the time. “Special interests now take precedence over the common good. …we now live in a diminished democracy …with ordinary citizens squeezed out of the public sphere by partisan ideologues and professional propagandists” (Dillard and Shen, 2013, p. 16). The Race to the Bottom is not dominant all the time, because loop dominance changes back and forth due to a variety of reasons beyond the scope of this version of the model.

Because the model is a simple system with few variables and loops, no formal loop dominance analysis is necessary (Oliva, 2016). Dominance is easily determined by experimentation, which agrees with inspection.

**Run 6. Reference mode 100 years starting in equilibrium.** This uses the same settings for false meme size, LTQ, and DTQ as run 5. Runs 6 and up start with the model in equilibrium and run for only 100 years. This magnifies graph changes over a short period of time. Because democracy is in crisis and rapid solution is urgently needed, we are most interested in what happens in the first twenty years.

**Run 7.** This and following runs begin pushing on the high leverage points to resolve the main root cause of low political truth literacy. In this run LTQ and DTQ are both raised from 20% to 50%. Optimal false meme size increases to 3.7. The result is percent rationalists soars from 39% to 75% and the Race to the Top becomes dominant. Virtuous politicians will tend to be elected and deceptive ones will not. Pushing on the high leverage points is beginning to have the intended effect.

However, 75% percent rationalists is not high enough. In a democracy the rights and desires of minorities must be respected and addressed. If 25% of voters prefer Race to the Bottom politicians, nations will be too distracted to focus efficiently on highly demanding problems, as seen in the disproportionate influence far right groups can have, such as the authoritarian populist wave in Europe of Le Pen in France, the Austrian Freedom Party in Austria, and the mis-named Center Party and Forum for Democracy parties in the Netherlands. In Germany, Denny (2021) found that the far-right Alternative for Germany (AfD), even though national polls show only 10% support, “poses a significant and complex threat to the German constitutional order. Highly organized and openly hostile to the rules binding other political actors, the German far right has outperformed its electoral support in shaping German society. In 2020, [one of Germany’s intelligence agencies] reported that the number of right-wing extremists in Germany has increased to 33,300, of whom 13,300 are thought to be willing to commit violence.” Historically, the far right is where authoritarianism support begins. We must do better.

**Run 8.** This tests the effect of raising LTQ and DTQ from 50% to 60%. Optimal false meme size increases to 3.9. Percent rationalists rises slightly from 75% to 83%. This is not enough, as discussed above by the destructive effect of Germany’s AfD party, which has only 10% support. We must therefore do even better.

**Run 9.** This raises LTQ and DTQ from 60% to 70%, and reduces optimal false meme size from 3.9 to 3.4. Citizens are getting so good at spotting political deception that degenerate politicians must scale down the size of their lies, so that smaller ones
slip through undetected. Percent rationalists rises a little more, from 83% to 89%. This is still not good enough, so let’s raise truth literacy again.

**Run 10.** LQT and DTQ are raised to 80%. Optimal false meme size falls to 3.1. Percent rationalists rises to 94%. As there is no empirical evidence to say if this is good enough, let’s raise it still higher.

**Run 11.** Finally, both high leverage points are raised to 90%. Optimal false meme size falls again to 1.9. Percent rationalists now reaches 97%, which appears sufficient. We hypothesize that a range of 80% to 90% LTQ and DTQ is required for satisfactory root cause resolution. Much further research is necessary, however, to determine the exact ranges for particular democratic systems, as well as how to measure LTQ and DTQ accurately.

Note how the 97% percent rationalists was not achieved by raising logical and democratic truth literacy (LQT and DTQ) to an unrealistically high level of 100%. A range of 80% to 90% appears achievable, considering the adult 2019 global literacy rate of 86% and the literacy rates of Central Europe and the Baltics (99%), East Asia and the Pacific (96%), and Latin America and the Caribbean (94%) (UNESCO, 2019).

**Comparison of the analysis to contemporary theories**

The analysis has been kept as simple as possible. It does not attempt to explain the rise and then the decline of democracy. It only explains why democracy is susceptible to decline, which is enough to solve the backsliding problem.

Fukuyama (2020) assigned three causes to democratic backsliding:

1. An ideological divide based on identity politics. The left represents oppressed minorities. The right “emphasizes a traditional kind of ethnically based national identity and worries that ‘our country’ is being taken over by a cabal of immigrants, foreign competitors, and elites who are complicit in the theft.”

2. Appearance of the global internet and social media. This has been used by the right, “by antidemocratic actors who have discovered that conspiracy stories and fabricated information often are rewarded with more clicks than the truth receives. … [Social media] permit like-minded individuals to find one another, not just in their own nations but around the world, while simultaneously shutting out criticism and disagreement.”

3. The decline in authority of trusted traditional social institutions for facts and news. “…the lethal threat to modern democracies is not the military coup, but rather a steady, gradual erosion of norms and institutions…” so that citizens no longer know what to trust, and are thus vulnerable to falsehoods.

The first cause, an ideological divide of the left and right, corresponds to the two dueling loops.

The second cause, use of the internet and social media by the right to amplify falsehoods, corresponds to a mechanism for exploiting the inherent advantage of The Race to the Bottom. Social media conversations and platform “news” algorithms can spread and amplify falsehoods in a vicious “echo chamber” feedback loop, unfettered by the
balancing influence of traditional media and the opinions of others with different and more truthful viewpoints. “...democracy [is] in crisis, buckling under the pressure of technological processes that [have] overwhelmed our collective capacity to tell truth from falsehood....” (Benkler et al., 2018, p. 4). Deliberately introduced “computational propaganda” and “manipulative disinformation campaigns” via social media channels by politicians, parties, and states is undermining democracy (Woolley and Howard, 2019, pp. 3–10).

The third cause, erosion of reliable sources of political information, is a standard strategy used by authoritarians to lower a population’s potential truth literacy. This erosion also occurred due to technology evolution, as over time the internet allowed (unreliable) social media to displace (reliable) traditional media.

Because formal causal chains and an appropriate modeling tool were not used, Fukuyama’s analysis lacks rigorous causal structure. His three causes lie on the superficial layer of the problem. Fukuyama was unable to penetrate to the fundamental layer.

Hannah Arendt’s classic work on The Origins of Totalitarianism addressed the fundamental question of describing how totalitarianism arises. She found three important preconditions: isolation, loneliness, and most importantly, susceptibility to propaganda:

The masses have to be won by propaganda. (Arendt, 1976, p. 39) … Before mass [movement] leaders seize the power to fit reality to their lies, their propaganda is marked by its extreme contempt for the facts as such, for in their opinion fact depends entirely on the power of the man who can fabricate it. (p48) … The ideal subject of totalitarian rule is not the convinced Nazi or the convinced Communist, but people for whom the distinction between fact and fiction (i.e., the reality of experience) and the distinction between true and false (i.e., the standards of thought) no longer exist. (p172)

Inability to distinguish between fact and fiction, true and false, is the same as low political truth literacy, which we found to be the main root cause of democratic backsliding. It should thus be no surprise that this is also the main root cause of susceptibility to totalitarianism. Once this root cause is resolved, the “ideal subject of totalitarian rule” would no longer exist in sufficient quantity for the path to totalitarianism to work.

The most useful theory we found was that of Benkler et al. (2018), not included in Waldner and Lust’s literature review. In an effort to understand the effect of propaganda on politics in a specific political system, the project analyzed four million messages in the US using their Media Cloud platform. America’s political spectrum has evolved into two opposing feedback loops, a right-wing “propaganda feedback loop” where politicians “compete on identity confirmation” regardless of the truth (p79), versus a centrist/left-wing “reality-check” loop that follows “institutionalized truth-seeking norms” where politicians “compete on truth quality and the scoop” (p77):

The American media ecosystem consists of two distinct, structurally different media ecosystems. One part is the right-wing, dominated by partisan media outlets that are densely interconnected and insular and anchored by Fox News and Breitbart. The other part spans the rest of the spectrum. It includes outlets from the left to historically center-right publications like the Wall Street Journal and
is anchored by media organizations on the center and center-left that adhere to [truth-seeking] professional standards of journalism. (p75)

[The media on the right uses] an echo chamber that radicalizes it inhabitants, destabilizes their ability to tell truth from fiction, and undermines their confidence in institutions. … The differences between the two media ecosystems are palpable. Despite extensive efforts, we were unable to find an example of disinformation or commercial clickbait started on the left, or aimed from abroad at the left, that took hold and became widely reported and believed…. By contrast, we found such instances repeatedly succeeding in the right-wing media ecosystem, with pervasive exposure and lasting effects on the [false] beliefs reported by listeners, readers, and viewers within that network. (p383)

Benkler’s findings are surprisingly close to what our own analysis found. The propaganda and reality-check loops correspond exactly to the Race to the Bottom and Top loops. This confirms that the Duelling Loops are firmly established in the US, the most influential democracy in the world. Ample evidence exists, e.g. (Cosentino, 2020; Heft et al., 2019; Pascale, 2019), that similar dynamics occur in many other democracies, particularly those with authoritarian tendencies.

Svolik (2019) found that “Electoral competition often confronts voters with a choice between two valid but potentially conflicting concerns: democratic principles and partisan interests.” These two choices correspond to the Race to the Bottom versus the Race to the Top. Why do voters choose Race to the Bottom candidates? Svolik argues the reason is political polarization: (italics added)

Each [politician] has succeeded in transforming his country’s latent social tensions into axes of acute political conflict and then presented his supporters with a choice: Vote for a more redistributive Venezuela, a migrant-free Hungary, a conservative Turkey—along with my increasingly authoritarian leadership—or vote for the opposition, which claims to be more democratic but offers less-appealing policies and leadership. […] This works because] ordinary people are willing to trade off democratic principles for partisan interests. [As a study showed.]

But from a root cause perspective, WHY is an authoritarian politician able to perform this transformation, where voters are clearly voting against their own best interests? By use of political deception to falsely polarize the electorate. WHY are they able to do that so easily? Because of the deeper cause of low political truth literacy. Polarization, a deceptive strategy, is an intermediate cause.

While all have useful concepts, none of the above theories or the six theory families examined by Waldner and Lust (2018) contain or attempt to find the problem’s formal causal structure. Instead, all are collections of related factors and plausible intermediate causes connected by problem stories.

The Truth Literacy Training study

Good “scientific modeling,” as Homer (1996) reminds us, requires testing of all key assumptions. “…every system dynamics model goes through an early exploratory stage in which some hypotheses are formulated with little or no empirical foundation. But
models that go no further than the exploratory stage should not be confused with those that are subjected to the rigors of scientific evaluation.” How much empirically based confidence can we have in the hypothesis that political truth literacy is low and can be raised to high in a practical manner, which forms the bedrock of the model?

The results reported in this paper are part of a larger body of work which designed a comprehensive collection of solution elements for pushing on the high leverage point of *raise political truth literacy from low to high*. As listed in Figure 1, nine sample solution elements were designed. The most promising one to develop first appears to be Truth Literacy Training, since it requires the least amount of work to develop, test, and implement for the highest impact.

**Solution element construction and study design**

Using our own software and database, a form of online Truth Literacy Training was developed using a long questionnaire supplemented by training materials (Figure 4). An elaborate series of pretests was used to refine training and testing design. At one point we found many subjects were not taking the training seriously, resulting in widely varying and mostly low scores.

The problem was solved by discovery of an insightful set of experiments (Sagarin et al., 2002) on “dispelling the illusions of invulnerability” to deceptive persuasion. The authors found it was not enough to inoculate subjects by exposure to deceptive statements and explanation of why they were deceptive. This failed to work because “our participants’ sense of unique invulnerability to deceptive ads left them unmotivated to use defenses against such ads.” This illusion of invulnerability caused subjects to believe they were not susceptible to deception, with the result that “they did not resist the ads containing illegitimate authorities [a form of deception] more effectively than did controls.” This was corrected by “demonstrating in an undeniable fashion that participants can be fooled by ads containing counterfeit authorities.”

To dramatically demonstrate to people that they are not invulnerable to deception, we changed the initial part of the training. After subjects in groups 2 and 3 answer questions for the first three statements (in the Review Section of Figure 4) and before any training has occurred, they read an educational item on: *The concept of truth literacy*. There they are shown their own answers (the first two are usually wrong) versus the correct answers. The item then says:

If you got all the answers right, congratulations. However, here’s how other people did. In a past survey with 34 participants, none got the answer to the first question right. Three got the answer to the second question right. On the third question 19 people got the answer right.

Why are the first two questions so hard? *It’s because they use clever forms of deception, which makes it terribly difficult to determine how true the claims are.*

The reason so many people got the third question right is it’s not deceptive. Generally, it’s much easier to spot the truth as opposed to deception, because we are so used to processing true statements from people we talk to, books we read, and so on.
This shocks people into realizing they are vulnerable to deception. From this point on, almost all take the questionnaire seriously. From the viewpoint of the elaboration likelihood model of persuasion (Dillard and Shen, 2013, pp. 137–149), elaboration motivation has increased from low to high. Subsequent training increases their elaboration ability, with the result that when training is complete, most deceptive persuasion attempts will be processed (elaborated) correctly and they will not be fooled.

Figure 4. Truth Literacy Training web page, group 3, claim and vote training. The subject has completed the training, done in the Getting Started and Review Section. They have just answered three questions concerning a statement about a Trade Agreement Treaty. All three answers are correct. The Personal Truth Test is shown on the right panel. Notes are the two vote training rules. Above The Personal Truth Test are summaries of the fallacies groups 2 and 3 were trained on. The subject scrolls to see all Reference Material. Using the left panel, subjects can navigate anywhere in the questionnaire to review their work. Checks indicate a completed item.
Using a Prolific online panel, the study was run on 93 US subjects randomly assigned to three groups. Demographics were age range 22 to 51, average age 31, 49% male. Educational levels were 34% high school, 55% college degree, 10% PhD. All were told this is a decision-making study for the purpose of improving the health of democracy.

**Group 1** (the control group) received training on the neutral topic of how democracy works.

**Group 2** received training on how to tell if a political claim (embedded in a political statement, such as the one in Figure 4) was true or false, by spotting the pattern of fallacy or non-fallacy used and using the Personal Truth Test, which includes the Strong Evidence Rule (Figure 5).

**Group 3** received the same training as group 2 plus training on how to vote correctly (given the perceived level of truth of a claim) by applying two rules: Reward the Truth Teller and Penalize the Deceiver. Total time for group 3 averaged 87 minutes, of which about one hour was training. Group 3 training involves 37 questions.

There is a 5-minute break after training for all groups, necessary to avoid fatigue and loss of interest on such a long questionnaire. A follow up study was run later using different statements.

In the test section of the questionnaire (called Decision Making Section in Figure 4), non-hot statements were presented in random order. Figure 4 shows how each statement is followed by three questions: (1) the truth question, (2) an open-ended question designed to maintain cognitive motivation and give us feedback, and (3) the vote question. The fictitious country of “Rutania” was used in statements to create interest and political realism without the bias a real country would have provoked.

**Definitions**

*Truth literacy* is the ability to tell truth from deception, to be able to “read” the truth.  

*Truth quotient* (TQ) is a measure of a person’s truth literacy in terms of their average ability to correctly process deceptive arguments in terms of how true an argument’s claim is, on a scale of zero to 100%. 100% is perfect truth literacy, which is not realistically possible due to the complexity and continual evolution of real-world deception. Like IQ or EQ, TQ measures an important aspect of intelligence. There are two components of TQ:

*Logical truth quotient* (LTQ) is the ability to logically tell if a deceptive claim is true or false. LTQ was measured by the percent correct for the truth questions for deceptive statements.

*Democratic truth quotient* (DTQ, aka political truth literacy) is the ability to vote correctly given a deceptive statement made by a politician. DTQ was measured by per-
cent correct for vote questions for deceptive statements. DTQ and LTQ are the two high leverage points on the Dueling Loops model (Figure 2).

Individual DTQ can theoretically never be higher than LTQ, since DTQ uses the results of LTQ as input. Study results support this prediction.

Deceptive statements contained six fallacies we found common in political appeals: cherry picking, ad hominem attack, appeal to emotion, strawman, false dilemma, and false fact lie, plus flawed application of the Strong Evidence Rule.

**General results**

Figure 6 summarizes study results. The results bring alive how the real world behind the system dynamics model behaves quantitatively.

Group 1 (the control group) corresponds to simulation runs 5 and 6, where LTQ and DTQ are both low. Group 2 claim training raises LTQ to high but has no effect on DTQ, which is still low. Thus group 2 also corresponds to runs 5 and 6, since any LTQ higher than DTQ has no effect on model behavior. Group 3 claim and vote training raises LTQ and DTQ to high, which corresponds to runs 10 and 11.

Figure 6 shows that LTQ is naturally low, at 8% for group 1. Voters not trained in logical truth literacy can spot a fallacy in a deceptive political statement an average of only 8% of the time. DTQ is also naturally low, at 2% for group 1. *Both are crucial findings and appear to explain why nations are so susceptible to a dominant Race to the Bottom and democratic backsliding.* While the study cannot say 8% and 2% are highly accurate measures due to the large confidence interval and low alpha, as well as the fact this was not a real-world test, the results indicate political truth literacy is low instead of medium or high in US voters. We expect this can be generalized to most political states, as some pretesting in other countries showed little difference, though our sample sizes were small.

---

*Figure 6. Results of the Truth Literacy Training study. Average scores and 95% confidence intervals for answers to deceptive statements are shown, with guessing levels and Cronbach’s alpha. Treatment groups were:

1 – Trained on neutral topic (control group)
2 – Trained on claims
3 – Trained on claims and vote*
The key data is LTQ and DTQ for groups 1 and 3. The large increases, from 8% to 76% and from 2% to 67%, 68-point and 65-point rises, suggest that Truth Literacy Training and other solution elements are capable of pushing on the high leverage point of 

raise political truth literacy from low to high 

successfully. Group 3 training averaged about one hour, indicating that Truth Literacy Training, such as in education systems and online training, will not require that much of a person’s time.

The follow up study 26 days later found LTQ and DTQ for group 3 had declined from 76% to 66% and 67% to 60%, 10-point and 7-point falls. After an average of 30 minutes of refresh training, LTQ and DTQ for group 3 rose to 75% and 70%, indicating regular refresh training of some type can work and will be required. Or it may be that like reading and writing literacy, once truth literacy matures and becomes the reasoning default and is exercised often enough, little decline will occur. Long term it may even rise, if reinforced by regular exposure to truth literacy practices in news media, such as a story centered on deceptive use of the cherry picking fallacy.

However, LTQ for group 1 was 22% and 20% for the pre-refresh and post-refresh training statements, versus 8% for the first questionnaire. This indicates spotting deception was substantially easier in the second questionnaire statements, and suggests there was more than the 10-point and 7-point declines noted above and that the refresh training may not have worked as well as the 75% and 70% indicated. The reason for this appears to be that the second questionnaire statements were developed after the first questionnaire was run. Without realizing it, we structured them slightly differently and frequently omitted stating how strongly supported the premises were. This caused the second set to be substantially easier than the first set, as it made fallacies easier to spot.

The problem is easily corrected. A more accurate measure of training persistence would require further statement testing/development and rerunning the study using balanced statements of equal difficulty in the first and second questionnaires. During this work the first and second questionnaires could be improved as needed.
Hypotheses results

In the first questionnaire, N=30, 30, and 33 for groups 1, 2, and 3. In the second questionnaire 26 days later, N=25, 27, and 24, an average dropout rate of 18%. A minimum sample size of 30 is required to assume results have a normal distribution and a reliable standard deviation can be calculated. Thus, the results of the first questionnaire are potentially reliable, while results for the second are slightly unreliable. It was our error to not use higher sample sizes for the first questionnaire to accommodate dropouts in the second questionnaire. Given the small sample sizes, we consider this a very preliminary study.

The study was designed to test six hypotheses. We reached these conclusions:

**H1. LTQ and DTQ can be accurately measured.**

Cronbach’s alpha is the standard measure of internal consistency, and hence reliability, of a set of Likert scale questions designed to measure a single construct (like LTQ or DTQ) in a single situation (such as groups 1, 2, and 3). The standard rules of thumb for alpha are > .9 Excellent, > .8 Good, > .7 Acceptable, > .6 Questionable, > .5 Poor, and < .5 Unacceptable (Gliem and Gliem, 2003).

The trend in the first and second questionnaires is that the more the training, the higher the alpha. The more a person is trained to make a discrimination, like spotting deception and deciding how to vote, the less they must guess when choosing an answer. Less guessing increases internal consistency, which increases alpha. The relatively high alphas for group 3 in the first questionnaire, .82 and .92, indicate the scale itself is sound. LTQ and DTQ can therefore be accurately measured for group 3, the trained population.

But what of group 1, the untrained population, which has alphas of .38 and .44 in the first questionnaire? Increasing the number of questions in a survey of this type (where alpha is high enough to be promising and under some situations is high, like group 3) can be expected to increase alpha to an acceptable level (Taber, 2018, p. 1287), as long as the new questions measure the same construct in a similar manner to the old questions. Thus, the results demonstrate that LTQ and DTQ can be accurately measured in general.

In the remaining hypotheses, only the 95% confidence intervals are considered.

**H2. LTQ and DTQ are currently low in the average voter.** This was well supported by the first questionnaire. Group 1 represents the average voter, who has never received the equivalent of Truth Literacy Training. Their LTQ and DTQ were low, 8% and 2%.

**H3. LTQ and DTQ can be raised to high via Truth Literacy Training.** This was also well supported by the first questionnaire. Because of training, LTQ rose from 8% to 77% and 76%. DTQ rose from 2% and 6% to 67%.

**H4. Truth Literacy Training on LTQ alone is insufficient to raise DTQ to above the minimum DTQ for a healthy sustainable democracy.** This was well supported by the first questionnaire. The average vote score for those receiving claim training alone was 6%. This shot up to 67% for the group receiving claim and vote training.

This was an astonishing discovery. Even if a person has been trained on how to tell whether a political claim is true or false, they are unable to translate the truth or falsity
of a claim into correct action. Instead, they choose all sorts of answers for the vote question. This indicates the average voter currently does not penalize deceptive politicians. Yet in a time when political deception is so rampant and the truth is so rare, why would anyone not want to strongly penalize deceivers? Why would anyone not want to strongly reward truth tellers? These two behaviors are required for democratic governments to work for the best interests of voters. We suspect the reason for this behavior is that hardly anyone has received the equivalent of Truth Literacy Training and in particular vote training, which is amazing simple. Vote training consists of following the two simple rules described in the lower right of Figure 4.

**H5.** *Training on LTQ and DTQ persists but falls over time.* This was partially supported by second questionnaire results.

**H6.** *The fall in LTQ and DTQ over time may be eliminated with sufficient refresh training.* This was weakly supported by second questionnaire results. However, we expect the problems causing weak results can be eliminated as discussed earlier.

In all cases, further research is required to more fully confirm, reject, or modify these hypotheses. Note that study results apply only to political statements using the fallacies subjects were trained on, and is too narrow to apply to real world behavior. A broader solution will require broader training and much further development.
Vote question results

<table>
<thead>
<tr>
<th>Treatment Groups</th>
<th>Deceptive statements</th>
<th>Non-deceptive statements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Answers should all be 9</td>
<td>Answers should all be 1</td>
</tr>
</tbody>
</table>

1. Trained on neutral topic (control group)

![Graph A](image)

2. Trained on claims

![Graph C](image)

3. Trained on claims and vote

![Graph E](image)

The Vote Question
If the election were held today and this was all the information you had, how much impact would what the politician claimed have on your decision to vote for or against the politician?

Answers
1. Very large increase in support.
2. Large increase in support.
3. Medium increase in support.
4. Small increase in support.
5. It would make no difference.
6. Small increase in opposition.
7. Medium increase in opposition.
8. Large increase in opposition.
9. Very large increase in opposition.

Figure 7. Distributions of the vote question answers for the first questionnaire. The correct answer is 9 for deceptive and 1 for non-deceptive statements. A normal distribution curve was added to chart A for discussion.

Let’s examine the vote question answers for the three groups in Figure 7:

**Group 1. Trained on neutral topic (control group)** – While the effect surely varies across political units and study samples, we hypothesize that the first row approximates how voters in most democracies behave today.

In chart A there’s more support than opposition in response to a deceptive political claim. This has not gone unnoticed by politicians willing to engage in deception. Notice how close the data comes to a normal distribution centered on the midpoint. This indicates a person’s level of political truth literacy is largely due to random factors (environmental and genetic chance) rather than the formal education seen in charts C and E.
Chart B shows untrained people lean toward supporting truth telling politician, but seem shy about supporting them strongly with answer 1. We offer no theory why this is so.

**Group 2. Trained on claims** – As discussed earlier in H4, the second row contains what to us is astounding counterintuitive data. Citizens trained on how to determine the truth of claims but not trained in how to vote correctly, intuitively lean in the correct direction on vote answers. *But very few choose the correct answers of 9 in chart C and 1 in chart D.* A surprising percentage (22% and 29%) chose answer 5, “It would make no difference.” That’s like saying “It doesn’t matter to me at all if a politician tells the truth or not.” *But yet it must, if democracy is to thrive.*

Similar observations apply to other incorrect answers. Choosing 4 and 6 is like saying “It barely matters to me whether a politician tells the truth or not.” Deviations from correct answers are why the vote training in group 3 is required.

Correct answer preference is worse in chart D than chart B. Claim training reduced ability to vote correctly. This too is puzzling behavior we cannot explain.

**Group 3. Trained on claims and vote** – The third row, if we could get enough voters there, would resolve the root cause by raising political truth literacy from low to high. For the solution element to work, we estimate only 5% to 15% of an electorate needs effective training since most elections are close. The biggest training impact would be on uncommitted young and swing voters. Voters already strongly committed to a false ideology will tend to resist change due to the deceptive power of motivated reasoning, as discussed later. Training is not urgently needed for voters already supporting truth-telling politicians. This suggests that initially, training should target those who would benefit the most. In the long-term, all citizens should be trained.

The training needs improvement to reduce confusion of some kind, indicated by the answer 1 spike in chart E and the answer 2, 5, and 9 spikes in chart F. These should all be near zero.

**Correct answers**

Political deception has become so destructive in today’s world that nothing less than answer 9 for deceptive statements and answer 1 for non-deceptive statements is required for democracy to thrive. Electing truth-telling leaders (in terms of the objective truth about what’s best to optimize the common good) must be the top priority of voters.

For deceptive statements, even small deviation from the correct answer matters, since that indicates a person has been partially deceived and that adds up. Koch and Arsendt (2017) surveyed the research on cumulative media effects and found consensus that “When media coverage of a certain aspect of reality is biased (i.e., differs systematically from reality), cumulative exposure to this coverage can bias recipients’ perceptions.”

Deception involves three components (Dillard and Shen, 2013, p. 37): (1) *Response shaping*, the initial formation of a new slightly different response to a stimulus, such as a slight amount of doubt climate change is real, a small dislike of immigrants, some doubt a politician not a liar, or skepticism to wearing face masks during the Covid19
pandemic. (2) Response reinforcement, where each repetition of a falsehood (often using different false evidence) strengthens the false belief response. (3) Response change, how a person responds to a stimulus based on the amount of false belief acquired.

The cumulative exposure effect occurs because of component 2. Each additional exposure reinforces the conditioned response. “The strength of people’s attitudes depends entirely on the number of incoming messages about the attitude issue they have processed” (p76).

Consider deceptive statements. What begins as a small error in reasoning, such as a small deviation from the correct answer, with many repetitions grows into a large error, such as reasoning that “Well, Politician A didn’t tell that big a lie. I’ll vote for him anyway.” Or they might reason “Just because Politician B denies climate change is real is not that strong a reason to oppose him. Other things matter more.”

For non-deceptive statements, small deviation from the correct answer also matters, since that indicates a person does not understand how to best support truth-telling politicians. These errors accumulate and can lead to large errors, such as: “I like the fact Politician C told the truth on such an important subject. But telling the truth is not that big a deal. Other factors like experience are just as important.”

Nullifying the deceptive power of motivated reasoning

The analysis found the high leverage point for resolving the main root cause of democratic backsliding to be raise political truth literacy from low to high. The study offers empirical evidence political truth literacy is presently low and can be fairly easily raised to high, though this was a laboratory study rather than an actual democratic system.

The cognitive mechanism employed in the training is high-speed pattern recognition, by spotting patterns of non-fallacies (truth) or fallacies (deception). Subjects were trained on how to spot one pattern of the truth: correct application of the Strong Evidence Rule. This inductive rule of logic appears to be the most common rule in true political statements. Subjects were also trained to spot flawed application of the Strong Evidence Rule plus six fallacies common in political falsehoods: cherry picking, ad hominem attack, appeal to emotion, strawman, false dilemma, and false fact lie.

This approach can nullify the deceptive power of motivated reasoning, a well-established theory explaining how biased decision-making works (Kunda, 1990). The theory explains why once a person is fooled into believing deceptive goals and facts, the person becomes highly partisan and their false beliefs are unshakable. Instead of thinking logically, they behave as The Rationalizing Voter, the title of Lodge and Taber’s magnum opus, (2013), which summarizes decades of empirical research. “In short, citizens are often partisan in their political information processing, motivated more by their desire to maintain prior beliefs and feelings than by their desire to make ‘accurate’ or otherwise optimal decisions” (p149). When a prior belief is false (such as non-whites are inferior or climate change denial), deception has occurred and partisan reasoning will be erroneous. The research centers on the John Q. Public simulation model (p28-
The purpose of deception is to create a false belief that benefits the deceiver at the expense of the deceived. Motivated reasoning theory divides false beliefs into two broad types: False goals and false facts used to rationalize false goals. In a democracy, false goals are those that do not benefit the common good (the majority of the electorate), and benefit special interests (a minority) instead.

False beliefs are created and strengthened by fallacious arguments, which work when someone fails to spot a fallacy. But once a person learns the spot-the-pattern of truth or deception technique of Truth Literacy Training, they are inoculated. Fallacies they have been trained on, or logic they cannot identify, can no longer be used to fool them into a false belief, so they never believe the false belief in the first place or may question a false belief already held. This depends on the level of truth literacy a person has attained.

The premise of motivated reasoning theory is that all reasoning is motivated to achieve either accuracy goals (slow thinking) or partisan goals (directional, fast thinking). With enough training and experience in how to “read” the truth by pattern recognition (claim training), and how to use that knowledge to act correctly (vote training), sufficiently correct accuracy reasoning can approach the speed of partisan reasoning and replace it, thereby becoming the reasoning default when important new political arguments or facts are encountered, or old ones need review.

“…the ways by which we are deceived are consistent and not so hard to recognize” (Jackson and Jamieson, 2007). A thorough program of pattern recognition training can create a reliable high-speed pattern recognition heuristic. With proper training, accuracy reasoning is now automatically used instead of partisan reasoning when confronted with new potentially false inputs, because accuracy reasoning is now fast instead of slow, and usually correct instead of so easily deceived. A person’s important political beliefs will now tend to be true instead of false, depending on their level of truth literacy.

Heuristics work by substituting fast and frugal reasoning for slower logical reasoning. “A heuristic is a strategy that ignores part of the information, with the goal of making decisions more quickly, frugally, and/or accurately than more complex methods” (Gigerenzer and Gaissmaier, 2011, p. 454). Pattern recognition allows irrelevant information to be ignored and not processed, which can speed reasoning so much it becomes instantaneous.

For example, the statement in Figure 4 contains 85 words about a Trade Agreement Treaty. As soon as a truth literate person spots the false dilemma pattern in the 6 words saying “There are only two choices here” and confirms there really are more than two choices (there almost always are in politics), they know the statement is fallacious and nothing else matters. Step 3 of the Personal Truth Test (Figure 4, right panel) applies.

Truth Literacy Training employs the preemptive aspect of inoculation theory. Innovating by training on logic pattern recognition instead of misinformation correction (such as fact-checks and news pointing out the truth), as we have done, Cook et al. (2017) found that inoculating subjects by training on spotting false balance and fake
expert strategies “neutralized” the negative influence of misinformation on perceived scientific consensus on climate change.

Our approach necessarily goes one step further by introducing vote training and greatly improves training effectiveness by adding a catalog of common fallacies and The Personal Truth Test. Drilling subjects on the catalog of common fallacies (and correct and flawed application of the Strong Evidence Rule) teaches high-speed pattern recognition in the same manner that students are drilled on letters of the alphabet, words, numbers, multiplication tables, object names from pictures of a cat, dog, etc. With sufficient practice citizens can now “read” the truth. Common fallacies can be spotted in seconds and false meme infection prevented most of the time.

Peering into the future, what might a society that has achieved universal truth literacy look like? The key cultural trait might be something like: (Boush et al., 2009, pp. 123–124)

Marketplace Deception Protection Skills…. A person who is skilled in deception protection will have well-learned mental procedures designed to detect, neutralize, resist, correct for, and penalize deception attempts…. More broadly, consumers adept at deception self-protection will learn to warn and protect friends, kin, and loved ones…. Most broadly, consumers must learn to adopt a deception protection goal as their default….

Discussion

Summary of findings

Our first research question was: (1) What is the key dynamic structure behind observed declines in liberal democracy and its main root causes?

Using system dynamics, The Dueling Loops of the Political Powerplace model of the democratic backsliding problem was constructed. The main root cause of the decline was found to be low political truth literacy. The high leverage point, where a small amount of solution force can have a large effect in resolving the root cause, is raise political truth literacy from low to high. The reason popular solutions have largely failed is they push on the intuitively attractive but low leverage points of more of the truth and misinformation correction.

Our second question was: (2) Can it be empirically demonstrated, on a preliminary basis, that the root causes exist and can be resolved?

In general, yes. Successful solution policies must focus on the high leverage point of raising political truth literacy from low to high enough to cause the Race to the Top to become the strongly dominant feedback loop. Since this is a novel claim and lies at the heart of the analysis, it was empirically tested with a single solution element in the Truth Literacy Training study. While the study has not been replicated and was a laboratory rather than a real-world study, results suggest that in the average democracy, political truth literacy is low and can be raised to high with a small amount of easily administered training. The training effect persists with a small decline, easily restored with
a short amount of refresh training. The training approach could be practically incorporated into a citizen’s normal education, which would now rest on a foundation of reading, writing, math, and truth literacy. While study results suggest Truth Literacy Training is a promising policy, we expect that many diverse solution elements will be required to fully and swiftly resolve the main root cause.

These findings provide the foundation of a comprehensive theory explaining democratic backsliding. Using a high-level causal diagram (Figure 1) paired with a system dynamics model (Figure 2), the Dueling Loops theory explains:

1. *Why democratic systems are backsliding.* This is due to the presence of the Dueling Loops structure and an unresolved main root cause (force R).

2. *Why popular solutions tend to fail.* Superficial solutions (force S) have pushed on intuitively attractive low leverage points.

3. *Why certain innovative solutions can be expected to succeed.* Fundamental solutions (force F) can push on the correct high leverage point in order to resolve the main root cause.

Building on these fresh insights, we offer conclusions and suggestions for practical next steps for applied research.

**Shifting from misinformation correction to misinformation prevention**

While fact-checks have greatly increased public awareness of the high level of deception in politics and the need to reduce the success of that deception, they have enjoyed less than the effectiveness expected. This is because fact-checks follow a strategy of misinformation correction rather than misinformation prevention.

Beginning in 2003 with FactCheck.org, the global rise of political fact-check organizations staffed by professional journalists has shined the light of truth on deception in politics. “Fact-check” has entered the popular lexicon. Articles, news hosts, and even presidential debate moderators routinely cite fact-checks. The fact-check movement has invented “a new style of political news, one that seeks to revitalize the ‘truth-seeking’ tradition in journalism by holding public figures to account for the things they say” (Graves, 2016, p. 6). In theory fact-checks should greatly reduce deception success.

However, in practice fact-checks and falsehood correction, the most popular solutions for reducing the damaging effect of political deception, are largely ineffective (Margolin et al., 2018; Oeldorf-Hirsch et al., 2020; Walter and Murphy, 2018) because they push on the low leverage point of misinformation correction, as diagrammed in Figure 1. “Once inaccurate beliefs are formed, they are remarkably difficult to eradicate. Even after people receive clear and credible corrections, misinformation continues to influence their reasoning: in cognitive psychology this is known as the continued influence effect of misinformation” (Swire and Ecker, 2018, italics in the original). The effect is so strong that “even if people do shift their opinion and acknowledge that information they previously believed to be true is incorrect, they are unlikely to change their voting preferences or feelings towards political candidates.” Others found the same result. “…journalistic fact-checks can reduce misperceptions but often have minimal ef-
fects on candidate evaluations or vote choice. … These findings suggest motivated reasoning can coexist with belief updating” (Nyhan et al., 2019). A recent meta-analysis of fact-checking studies (k = 30, N = 20,963) concluded that “In line with motivated reasoning literature, the effects of fact-checking on beliefs are quite weak and gradually become negligible the more the study design resembles a real-world scenario of exposure to fact-checking” (Walter et al., 2020).

Two further reasons fact-checks are ineffective are that fact-checkers have never solved the problems of:

1. How to get people to read a fact-check corresponding to a particular falsehood.
   “We almost never observe respondents reading a fact-check of a specific claim in a fake news article that they read” (Guess et al., 2018). Misinformation cannot be corrected if people never read the correction.

2. How to fact-check most potential falsehoods. Currently this is prohibitively expensive, though this is expected to improve in the long-term with use of artificial intelligence and natural language processing (Graves, 2018).

Both problems can be solved by teaching citizens how to be their own truth-checkers with approaches like Truth Literacy Training. “Give a man a fish and you feed him for a day. Teach a man to fish and you feed him for a lifetime.” The training goal is to become your own truth-checker rather than fact-checker, because the burden of fact-checking is not on the consumer of a claim but on the creator. Claim authors, such as politicians and writers, must cite reliable sources for all important premises (facts) that are not common knowledge. If this information is not provided, The Personal Truth Test teaches citizens to reject the claim since its truth is unknown and quite possibly false.

As of June 2021 there were 341 fact-checking organizations in 102 countries (Stencel and Luther, 2021). Given the ineffectiveness of fact-checking, the analysis results, and a modest amount of empirical evidence that political truth literacy is low and can be raised to high with Truth Literacy Training, it would behoove fact-checkers, journalists, and news organizations to shift from misinformation correction to misinformation prevention, by switching to the high leverage point of raise political truth literacy from low to high. “An ounce of prevention is worth a pound of cure” (Franklin, 1734, referring to the need to protect towns from fires by various prevention practices).

**Averting disaster in the nation that invented modern democracy**

Democratic backsliding has reach crisis levels in the US. A recent poll found that:

… large numbers of Joe Biden and Donald Trump voters view the other party with fear and contempt. …supermajorities of voters in each camp believe the other side is bent on destroying the country. More than 80 percent of Biden and Trump voters agree that elected officials of the other party “present a clear and present danger to American democracy.” More than 70 percent of both sets of voters believe that some extreme media voices on the other side should be censored “despite the U.S. Constitution’s First Amendment.” More than 75 percent
of Biden and Trump voters believe that Americans who strongly support the opposite party also threaten the American way of life. In short, politics has stopped being about how to govern a shared country and is more about a naked, “Lord of the Flies”-style struggle for power. (Olsen, 2021)

This is a recipe for disaster. “The United States is heading into its greatest political and constitutional crisis since the Civil War, with a reasonable chance over the next three to four years of incidents of mass violence, a breakdown of federal authority, and the division of the country into warring red and blue enclaves” (Kagan, 2021). “[In America] the old labels ‘left’ versus ‘right’ are fast becoming outdated. Today, it’s democracy versus authoritarianism” (Reich, 2021).

The high leverage point for resolving the main root cause of democratic backsliding is raise political truth literacy from low to high. Can high leverage point solutions be devised and implemented fast enough to avert the US tipping into an authoritarian state? We see this as realistically possible, considering how close US presidential and other elections have been recently, some evidence that pushing on the high leverage point can work (such as our study, Cook et al. (2017), and Sagarin et al., (2002)), and the enticing fact that the high leverage point has never been pushed on before with large-scale solution elements.

There are numerous ways to push on the high leverage point in addition to Truth Literacy Training. Figure 1 lists eight additional sample solution elements. Many more are possible. The training itself would need a much broader catalog of fallacious patterns to spot, like conspiracy theories, false populism, pushing particular emotional hot buttons (especially fear, anger/rage, and hate), sophisticated misinformation campaign patterns like deception by omission and saturation (Kopp, 2006), and the many custom new forms of deception that constantly appear. The training can be woven into all forms of journalistic discourse by elements like centering a news story on the fallacy(s) used, the importance each person’s DTQ plays in the health of democracy, maxims like “Don’t be fooled” and “Don’t be a pawn in someone else’s game,” the favored forms of deception for certain politicians and parties, the four characteristics of an authoritarian to watch for (Levitsky and Ziblatt, 2018, listed on pp 21-22), and so on. Just as the medical profession can inoculate a population against a physical virus to stop a pandemic, so too can education and news professionals inoculate a population against the mental virus of political deception to stop democratic backsliding.

If a high leverage point strategy works, the increase of Race to the Top votes should, we hope, be enough to offset the rapidly increasing attempts by US Republicans to subvert upcoming elections. Hasen (2022) lays out in stark detail how a massive election subversion attack is underway, by use of voter suppression, voting fraud, suppressive election administration, violent acts, and even overturning of undesired election results. “The time to act is now, before American democracy disappears.” Hansen argues the attack can be stopped two ways, by legal changes related to election reform and “political organization… to reinforce norms respecting the rule of law and fair election processes.”
The high leverage point strategy offers a third way. We expect it is likely the only viable way, since Republicans are already blocking meaningful election reform and have demonstrated the ability to destroy rule of law and election norms among a sufficient percentage of voters with ease. Hansen’s two ways, legal changes and reinforcing norms, are forms of more of the truth (Figure 1) about what people should do and are thus low leverage point strategies. Republicans have proven that the right political deception will cause people to ignore the law and important social norms, notably with the Big Lie that Trump actually won the 2020 election and the January 6, 2021 attempted coup. Superficial solutions based on more of the truth have done little to prevent deception success. But fundamental solutions like Truth Literacy Training can, because they strike at the root of the problem.

Considering how permanent the transition to authoritarianism tends to be among superpowers (such as Russia from Lenin to Putin beginning in 1922 and communist China beginning in 1949), the importance of moving quickly here cannot be understated.

Scientific management of the key requirement for a healthy democracy

The analysis behind the Truth Literacy Training study implies a hypothesis the study itself cannot measure or test. Given the indispensable role of the voter feedback loop in modern democracy, we propose what can be called the minimum political truth literacy requirement of democracy: A certain average minimum democratic truth quotient (minimum DTQ) is required for a healthy sustainable democracy, defined as one that can consistently achieve its top common good governance goals indefinitely.

Above this minimum, the dominant loop in a democratic system becomes The Race to the Top Among Politicians and that democracy works as intended. On the other hand, the further a democracy falls below this minimum, the more The Race to the Bottom dominates. As a consequence, the more susceptible that democracy is to backsliding and the lower the benefits it will tend to deliver to the majority of its citizens, since the more it backslides, the greater the share of benefits delivered to the minority (powerful special interests, aka the ruling elite) and the less priority given to top common good problems, such as war, high economic inequality, environmental sustainability, and in particular the looming climate change crisis.

We therefore argue that once a democracy is established, the key requirement for its health is the electorate’s average DTQ. If this is sufficiently high, the system automatically improves all other factors as needed, because the system self-manages itself in the preferred direction of optimizing the common good of all, using The Race to the Top feedback loop. These other factors include per capita income, inequality, political institutions, and political culture, all seen to be crucial for democratic consolidation (Dahl and Shapiro, 2015, p. 196). The better the democratic system managers elected, the better factors like these will be managed. A dominant Race to the Top is not just an accountability loop. Managed properly, it is an optimization loop.

If more model detail was added to Figure 2 it would show the Race to the Top acting together with a Self-Management of the Common Good of All goal-seeking loop, whose explicit goal is optimizing the long-term common good. The goal causes politi-
cians to cooperatively compete to see who can attract the most supporters by telling them the objective truth about what’s best for the general welfare of all. To maintain symmetry there would be an identical opposing loop added to the Race to the Bottom. Its goal would be the short-term maximization of my competitive advantage (such as wealth and power), which is what powerful special interests strive for, due to the endless struggle of survival of the fittest. This goal-seeking structure provides (at the root cause level) the foundation for the mechanism required to achieve the social homeostasis and automatic regulation so essential to the healthy and successful life of an organism, whether it be a biological species or a social system (Richardson, 1991, p. 49 and 52).

The self-management process centers on scientific management of the electorate’s average DTQ, by measurement and improvement as necessary, based upon the theory of democratic backsliding embodied in Figures 1 and 2. Comparable self-management analogies are: (1) The way the business world centers on scientific management of short-term profit maximization, based on the theory of double entry accounting and a long cumulative history of how different practices and forms of management structure work best in different business situations, and (2) The way science itself centers on the scientific method and the goal of steady accumulation of useful cause-and-effect knowledge, based on the theory of controlled experimentation and knowledge of past experimental results.

Here the three limiting constraints to be maximized/optimized are DTQ, profits, and quality of the scientific method. Once a system’s main constraint (aka bottleneck or limiting factor) is fixed by resolving its root causes, others will appear. For more on this viewpoint, see the literature on Theory of Constraints (Watson et al., 2007).

What the minimum DTQ requirement is for different political systems, how DTQ behaves dynamically in real systems, how DTQ and the other factors it controls can best be effectively self-managed, and what other root cause forces require management offers a rich area for further political science research.

**Beyond the exploratory model stage**

We are under no illusion that the exploratory system dynamics model presented in this paper is sufficient to lead to solution of the democratic backsliding problem in the fastest, most efficient manner possible. But we do see the model as a productive first step in a new direction, as an archetype at the heart of political system behavior. What’s needed next “is not necessarily a large model, but a model which takes into account a wide range of known details and which is therefore capable of making predictions with levels of confidence and insight [much] greater than those of an exploratory model” (Homer, 1996).

Where is the present model weak or flawed? How can the model be revised and calibrated to match the backsliding behavior of particular nations, regions, or the planet as a whole? What can explain the three waves of democratization and autocratization (Lührmann and Lindberg, 2019) and their timing? Why have the second and third (intermediate) causes identified by Fukuyama (2020) in this paper’s section on “Compari-
son of the analysis to contemporary theories” had such a profound effect? Since difficult complex system problems tend to have multiple root causes, what other root causes exist? Most importantly, how can solution be accelerated? Answers to questions like these are urgently needed, and offer a rich area for further system dynamics research.

Acknowledgements
(To be added later.)

Supporting information
The simulation model and study data spreadsheet may be found in the online version of this article at the publisher’s website. For examination of the study online questionnaire and free use of the software, please contact the authors.

References


Reich, R., 2021. What’s really wrong with the mainstream media. Guard.


