Democracy's eternal vulnerability: Increasing resilience to disinformation by raising the two components of political truth literacy

1. Method of analysis

1.1. *Why we selected root cause analysis (RCA)*

Current theory cannot comprehensively explain the causes of democratic system susceptibility to disinformation, indicating that the complexity of the backsliding problem exceeds the present capacity of traditional political science methods. The same holds for other members of the class of difficult large-scale social problems, including war, environmental sustainability (including climate change), high economic inequality, large economic recessions, systemic discrimination, government corruption, failed states, and urban decay. Howlett (2009, italics added) describes this as the problem of low "*policy analytical capacity* [which prevents] effectively dealing with many complex contemporary policy challenges."

The difficulty of this class of problems arises from *high dynamic complexity*, which stems from the many interactions among social agents over time. Sterman (2000, p. 22) describes how these problems are *constantly evolving, tightly coupled, governed by feedback loops, nonlinear, history-dependent, self-organizing, adaptive, and highly counterintuitive*. Such characteristics make these problems very difficult or impossible to solve, unless their essential causal structure (ECS) is identified. Methods suitable for this class of problems must therefore have the ability to identify ECS as a central feature in order to construct a *glass box model*, "which overtly displays all [relevant] mechanisms and functions of the system being modeled" as opposed to a *black box model*, where "the computations are hidden and relationships between the variables of the system can only be inferred" (Blumschein et al., 2019, p. 9).

Root cause analysis (RCA) is, by definition, the only known method for reliably solving difficult complex causal problems. A *causal problem* occurs when problem symptoms have causes. Thus, all causal problems arise from their root causes. *RCA* is the systematic practice of finding, resolving, and preventing the recurrence of the root causes of causal problems. Reliability is an emergent property of the structured, systematic nature of the process. Consequently, any other method capable of reliably solving difficult complex causal problem must include RCA's core features, and is thus a superset of RCA. This is analogous to the way the scientific method is, by definition, the only known method for producing reliable new cause-and-effect knowledge. Countless fields implicitly employ RCA without calling it RCA, such as the way doctors are training to diagnose and treat the root cause(s) of illness.

The analytical power of explicit RCA may appear novel (even paradigm-shifting) to social scientists, who are generally not trained in RCA, as evident in standard textbooks such as (Johnson et al., 2019; King et al., 1994; McNabb, 2021; Singleton & Straits, 2005). By contrast, in the business world the RCA paradigm has long been taught and applied, and the terms "root cause" and "root cause analysis" are used routinely. According to the Harvard Business School, "Root cause analysis is a critical organizational leadership skill for effectively addressing problems and driving change" (Han, 2024).

Other analytical methods commonly used in the social sciences, such as comparative analysis, statistical analysis, path analysis, process tracing, agent-based modeling, expert opinion, case studies, and trial and error via experimentation (hill climbing), perform well on less complex problems. But when faced with high dynamic complexity, these methods lack the required *analytical capacity* because they are not designed to identify ECS and use that knowledge to generate *glass box models* and causally-informed solution strategy hypotheses, *a task RCA is* explicitly *designed to do with ease* (Anderson & Fagerhaug, 2006; Doggett, 2004; Okes, 2019). RCA is a high-level process that uses hundreds of specialized sub-tools (George et al., 2004, 100 tools; Pyzdek, 2003, over 100 tools; Tague, 2023, 142 tools) as needed, including statistics and experimentation.

For these reasons we selected RCA as the method of analysis.

1.2. How RCA works

RCA works by starting at problem symptoms and asking "WHY does this occur?" until the root causes are found, a generic procedure known as the *Five Whys* (Liker, 2004, pp. 252–256; Ohno, 1988, pp. 17–18). 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 with practical solutions, without side effects that create other equal or larger problems. *Resolved* means the problem will probably not recur due to that root cause.

A *causal problem* occurs when problem symptoms have causes, such as illness or a car that won't start. By contrast, non-causal problems include information search problems, math problems, scientific discovery problems, and puzzle solving. The golden rule of RCA is: *All causal problems arise from their root causes.*

The basic RCA process is generic and must be wrapped in a process suitable for a particular problem class. Examples of widely used wrapper processes for high-complexity problems are:

1. Six Sigma for process control (Pyzdek, 2003).
2. Lean for manufacturing (Womack et al., 1990).
3. Fault tree analysis for system failure incidents (Yazdi et al., 2023).
4. MECE issue trees (Chevallier, 2016) for particularly difficult business management problems.

RCA-based processes have proven so effective that Six Sigma is used by 100% of aerospace, motor vehicle, electronics, and pharmaceutical companies in the Fortune 500 and 82% of all companies in the Fortune 100 (Marx, 2007). Lean is the global best practice for large-scale manufacturing (Nguyen, 2018). MECE issue trees are the core tool used by the world's top three business management consultancies: McKinsey, Bain, and BCG.

While RCA is beginning to make inroads into social science, such as Wagner's (2014) eight-step RCA process for public policy pedagogy, no RCA-based method suitable for difficult large-scale social problems was found. We were thus compelled to develop one, a common occurrence with novel problem types. For example, "After extensive review, NASA found that none of the commercially available tools and methods would support a comprehensive root cause analysis of all the unique problems and environments NASA faces," causing NASA to create RCAT, their own Root Cause Analysis Tool (Uusitalo, 2011, p. 73). Similarly, our research project produced *social force diagrams*, a fill-in-the-blanks template that guides RCA.

RCA is not a panacea. However, when thoughtfully applied to difficult complex causal problems, it can lead to deep insights and fundamental solutions that no other problem-solving method can find reliably. For this reason, RCA has become *the* leading tool in the business world for difficult complex causal problems.

1.3. Limitations

The history of industrial RCA shows the limitations of the method depend primarily on the suitability of the RCA-based process used for a particular class of problems and the maturity of that process, including training program quality and training levels of process users. Any limitations encountered can eventually be over-come (unless that is physically impossible) by continuously improving the process. Over many decades, results that were previously considered beyond reach become routine. The archetypical example is the rise of Japanese global motor vehicle production market share, from 2% in 1955 to 28% in 1981, due to superior quality and lower price (Womack et al., 1990, p. 69). Part of the post-war Japanese economic miracle, the critical catalyst was an RCA-based process (the Toyota Production System, later standardized as lean) invented by Toyota and shared across Japan beginning in the early 1950s.

A physical limitation of the simulation models sometimes used in causal analysis is that due to sensitive starting conditions, accurate long-term predictions of large social systems and some types of natural systems is impossible. Examples are weather, stock markets, recessions, and wars. An accurate prediction problem is an information search problem, not a causal problem, since one is searching for initial starting conditions and system structure. The simulation model built for the analysis does not have this limitation, since it's a non-calibrated concept model with very simple behavior. A complex calibrated model would have some of this limitation.

Another physical limitation is the system must be inspectable. If portions of the system are kept secret, they cannot be inspected to identify causal structure. This applies particularly to social systems. Human minds cannot be easily inspected to find causal structure leading to behavior, so they largely remain a black box. However, human behavior traits like intelligence and truth literacy can be measured and improved without deeper knowledge of their causal structure. If a political system keeps portions of its behavior secret, as most authoritarian states do, that system cannot be fully modeled since it's not fully inspectable. However, what matters can usually be modeled by inspection of visible behavior.

1.4. Social force diagrams

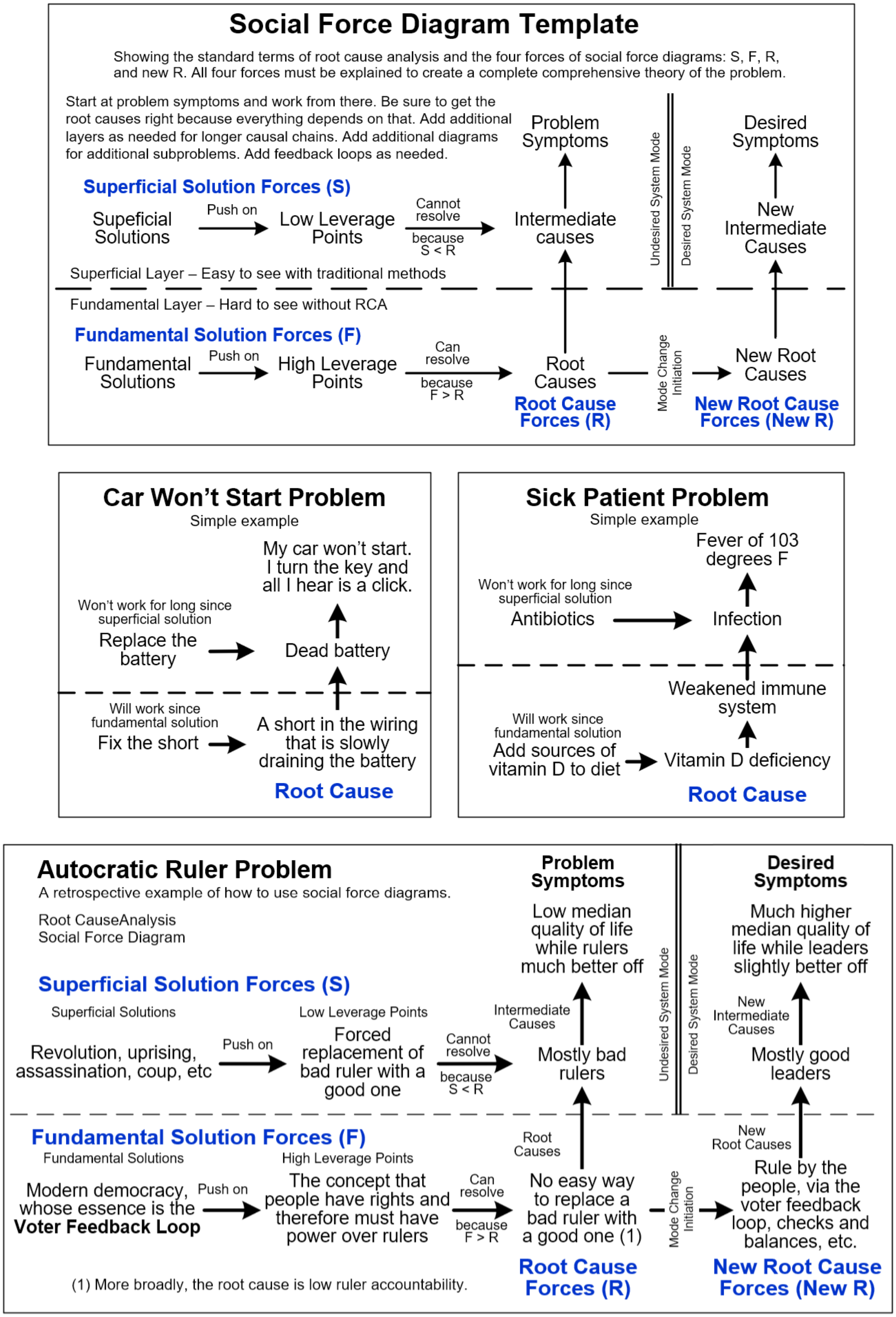

Figure 1. Social force diagram template, two simple examples, and a retrospective example.

Social force diagrams (Figure 1) are organized into two main layers:

1. The **superficial layer** of the problem, where intermediate causes are so easy to see they are erroneously assumed to be root causes, and
2. The deeper **fundamental layer**, where by understanding the problem's deeper structure its well-hidden root causes may be found. Without analysis of the fundamental layer, difficult problems tend to remain stuck in the superficial layer for a long time or indefinitely.

The two layers allow avoidance of the *Superficial Solutions Trap*. The trap occurs when problem solvers unknowingly assume intermediate causes are root causes, and then develop solutions based on that false assumption. This leads to solutions directed toward intermediate causes rather than root causes. Superficial solutions can *never* resolve root causes because root cause forces exert a greater force on intermediate causes than superficial solutions, regardless of how well funded, managed, or promoted those solutions are. Grainger (2025, p. ix) describes how without RCA, "organizations repeatedly solve the wrong problems. Without understanding the real root causes, teams end up treating symptoms rather than fixing the underlying issue." All the superficial solutions in Figure 2 (in the main paper) fell into the Superficial Solutions trap.

To demonstrate how social force diagrams work, two simple examples are shown.

Also shown is a retrospective example: the Autocratic Ruler Problem. One of history's most intractable problems was autocratic rule by countless warlords, dictators, and kings. The chief symptom was *low median quality of life while rulers are much better off*. WHY did that occur? Because of *mostly bad rulers*, who concentrated income, wealth, and power into the ruling elite.

The historic leverage point was *forced replacement of a bad ruler with a good one*, using solutions like *revolution, uprising, assassination, coup, etc*. While these solutions offered short-term gains, they failed to result in permanent long-term change because new bad rulers appeared. This repeated solution failure indicates *mostly bad rulers* was an intermediate cause. If the problem is solvable, there must be a deeper cause.

After thousands of years of painful trial and error, problem solvers intuitively found the deeper root cause. It was *no easy way to replace a bad ruler with a good one*. What would justify this easy replacement? It was *the concept that people have rights, and therefore must have power over their rulers*. This, in the hands of the writers of the Magna Carta and much later the US and French constitutions, crystallized the solution of *modern democracy, whose essence is the Voter Feedback Loop*. Once the first few countries adopted the solution, the benefits were so attractive that a feedback-loop-driven systemic mode change occurred, and democracy swept across much of the world.

1.5. The five steps of applying social force diagrams

1. Formal problem definition in terms of present symptoms, desired symptoms, and constraints.
2. Using RCA, identify superficial layer structure in order to learn from the past.
3. Using RCA, identify fundamental layer structure with emphasis on getting the root causes right.
4. Develop solution element candidates and test them on high leverage points.

5. Iterate until a small group of solution candidates have a high probability of success, as determined by testing as necessary depending on requirements, which may include real-world testing. These become solution options.

Note that our study's requirements were only to test H1 and H2 as stated in the main paper, on a proof-of-concept basis. Thus, there was no requirement for real-world testing or to iterate until done. Our study only performed steps 1 to 3, and step 4 for a single solution candidate, Truth Literacy Training. Steps 4 and 5 allow the method to scale up to large, real-world applications.

1.6. The four forces of social force diagrams

Social force diagrams focus on understanding four key forces: S, F, R, and new R. Superficial solutions (force S) fail because *force S is always less than root cause forces* (force R), indicated by **S < R**. By contrast, if fundamental solutions (force F) are properly designed (especially their impact on feedback loop structure), *force F can exceed force R*, indicated by **F > R**. This leads to a systemic mode change, during which the old R is replaced by a new R, and the problem is solved. The new R must be engineered to be strong enough and self-managing enough to permanently hold the system in the solved mode, due to the way force F fundamentally changes critical feedback loop structure and loop dominance.

If analysis shows no F > R exists (no resolvable root cause is found), the problem is unsolvable as defined. In this case problem definition (problem symptoms) can sometimes be relaxed to make the problem solvable, such as raising the maximum allowable global temperature rise for climate change to make that problem solvable. Otherwise, the problem should be declared unsolvable.

A *low leverage point* is a specific place in a system's feedback loop structure that, if solutions elements are applied, will not lead to solution, due to S < R. By contrast, a *high leverage point* is a specific place in a system's feedback loop structure that solution elements can push on in order to efficiently resolve the connected root cause, since F > R. A high leverage point is thus a high-level solution strategy.

Once all four forces are understood and key assumptions have been tested by measurement or experimentation, the analyst has a sufficiently complete scientific theory of the problem. Each force provides an explanatory tenet of the theory. This gives *the four requirements for a comprehensive theory of a difficult large-scale social problem.* The theory must identify the four forces and explain them in this manner:

1. **Social Force S** – Why past solutions failed (because S < R).
2. **Social Force R** – Why the problem occurs (because R is unresolved).
3. **Social Force F** – Why fundamental solutions can be expected to succeed in causing the desired mode change (because F > R).
4. **New Social Force R** – Why the mode change will be relatively permanent (because the new R contains self-sustaining feedback loops).

Due to problem complexity, different analysts will produce different diagrams/models and will often reach different policy conclusions. This is the well-known problem of model uncertainty/variation (Young, 2009). The standard solution is to specify requirements and use them for evaluation, a process step especially important for complex problems (Höge et al., 2018). The above four requirements provide this for the class of difficult large-scale social problems.

References

Anderson, B., & Fagerhaug, T. (2006). *Root Cause Analysis: Simplified Tools and Techniques, Second Edition*. ASQ Quality Press.

Blumschein, P., Hung, W., Jonassen, D., & Strobel, J. (2019). *Model-Based Approaches to Learning: Using Systems Models and Simulations to Improve Understanding and Problem Solving in Complex Domains*. Sense Publishers.

Chevallier, A. (2016). *Strategic Thinking in Complex Problem Solving*. Oxford University Press.

Doggett, A. M. (2004). A Statistical Comparison of Three Root Cause Analysis Tools. *Journal of Industrial Psychology*, *20*.

George, M., Maxey, J., Rowlands, D., & Price, M. (2004). *The Lean Six Sigma Pocket Toolbook: A Quick Reference Guide to 100 Tools for Improving Quality and Speed*. McGraw Hill.

Grainger, A. (2025). *Root Cause Analysis: A Leader's Guide*. Grainger.

Han, E. (2024, December 3). Root cause analysis: What it is and how to perform one. *Harvard Business School - Business Insights*. https://online.hbs.edu/blog/post/root-cause-analysis

Höge, M., Wöhling, T., & Nowak, W. (2018). A Primer for Model Selection: The Decisive Role of Model Complexity. *Water Resources Research*, *54*(3), 1688–1715. https://doi.org/10.1002/2017WR021902

Howlett, M. (2009). Policy analytical capacity and evidence-based policy-making: Lessons from Canada. *Canadian Public Administration*.

Johnson, J., Reynolds, H. T., & Mycoff, J. (2019). *Political Science Research Methods, Ninth Edition*. CQ Press.

King, G., Keohane, R., & Verba, S. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

Liker, J. (2004). *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer*. McGraw Hill.

Marx, M. (2007). Six Sigma Saves a Fortune. *Six Sigma Magazine*.

McNabb, D. (2021). *Research Methods for Political Science: Quantitative, Qualitative, and Mixed Methods Approaches, Third Edition*. Routledge. http://www.lib.ysu.am/disciplines_bk/5011976daf5b5916794c6966cf0b02b3.pdf

Nguyen, H. (2018). Is Lean Manufacturing Still Relevant Today? *TRG International*. https://blog.trginternational.com/is-lean-manufacturing-still-relevant-today

Ohno, T. (1988). *Toyota Production System: Beyond Large-Scale Production*. Productivity Press.

Okes, D. (2019). *Root Cause Analysis: The Core of Problem Solving and Corrective Action*. ASQ Quality Press.

Pyzdek, T. (2003). *The Six Sigma Handbook: A Complete Guide for Green Belts, Black Belts, and Managers at All Levels*. McGraw Hill.

Singleton, R., & Straits, B. (2005). *Approaches to Social Research, Fourth Edition*. Oxford University Press.

Sterman, J. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex*

World. Irwin McGraw-Hill.

Tague, N. (2023). *The Quality Toolbox*. Quality Press. https://www.google.com/books/edition/The_Quality_Toolbox/0hLqEAAAQBAJ?gbpv=1

Uusitalo, I. (2011). *Review of security testing tools*. Diamonds Consortium. https://publica-rest.fraunhofer.de/server/api/core/bitstreams/3062e987-a93a-420e-8cf1-e7679835aff8/content

Wagner, T. (2014). Using Root Cause Analysis in Public Policy Pedagogy. *Journal of Public Affairs Education*.

Womack, J., Jones, D., & Roos, D. (1990). *The Machine that Changed the World: The Story of Lean Production*. Harper Collins.

Yazdi, M., Huang, H.-Z., Adumene, S., Mohammadpour, J., Zarei, E., Li, H., & Pirbalouti, R. (2023). Fault tree analysis improvements: A bibliometric analysis and literature review. *Quality and Reliability Engineering International*. https://onlinelibrary.wiley.com/doi/pdf/10.1002/qre.3271

Young, C. (2009). Model Uncertainty in Sociological Research: An Application to Religion and Economic Growth. *American Sociological Review*, *74*(3), 380–397. https://doi.org/10.1177/000312240907400303