Disinformation and Trust in Media and Institutions

Matteo F. N. Giglioli[1]

Abstract: There is an intuitive plausibility to the link between disinformation and trust in media and institutions. Although social theory has investigated the concept of trust extensively, there are methodological problems (related to logical consistency, latency, salience, and aggregation) in applying it directly to causal accounts of speech acts such as disinformation. Nonetheless, the presence of an overarching empirical phenomenon, the contemporary crisis of trust in institutions, shapes the social context in which disinformation occurs in advanced industrial societies. Disinformation can be conceptualized as a strategy to exploit such a state of affairs. An understanding of the different pathways these exploits can take is beneficial in calibrating policy interventions, but typically modifying social trust is not a lever directly available to policymaking. Hence, the trust deficit is best seen as a ceiling to the effectiveness of disinformation-fighting endeavors.

This chapter endeavors to do two things. In the first part (sections 1-3), the concept of social trust is explored, the main contemporary empirical finding (the generalized crisis of trust) is presented, and the differences in scope and intentionality between trust and disinformation are highlighted. Building on these notions, in the second part (sections 4-5), the potential causal links between trust and disinformation are detailed, with a view to guiding further empirical exploration and optimizing policymaking interventions.

1 Social Trust

Trust is understood to be a basic mechanism allowing for the development of social cooperation, hence the deployment of social power and the growth of human civilization (Mann 2012). Evidence of the importance of trust can be found in many places, ranging from long-term, macro effects of trust differentials on human flourishing (Banfield & Banfield 1958, Putnam et al. 1992, Fukuyama 1996, Tilly 2005) to the contemporary micro behavior of corporate actors, as displayed in Public Relations (PR) budgets, reputation management, communication consultants, and so forth. In particular, some measure of diffuse trust is generally considered indispensable for liberal democratic societies, as it underpins the existence of the public sphere and makes possible the open exchange of views leading to the development of public opinion and the steering of collective decision-making.

From a theoretical point of view, trust is typically conceptualized by analogy to an archetypal case of a single individual forming a belief about a statement, behavior, or intention of another. Therefore, there are clearly psychological (and, in some cases, ethological) micro-foundations that are relevant to the concept (Krueger & Meyer-Lindenberg 2019). However, for the purposes of a discussion of its relation with disinformation, trust is more often conceived of in aggregate terms. In other words, what

[1]M.F.N. Giglioli
NYU Florence
Via Bolognese 120
Firenze 50139
Italy
matteo.giglioli@nyu.edu

matters typically is group beliefs, about patterns of statements, behaviors, or intentions, producing reputations and affecting the trustworthiness of corporate bodies.

In particular, when we consider the information ecosystem, questions of trust or its absence may arise with reference to participants (speakers) in a debate, the subjects on which they speak, the medium or intermediaries through which they communicate, and the networks of sociability within which the listeners are embedded and which shape their assessments of debates and emerging beliefs. Furthermore, in most of the cases of interest to the topic of disinformation, what is at stake is what might be called 'agonistic' trust, namely the presence of multiple incompatible claims of fact calling for a choice between them.

The concept of trust is easy to understand intuitively, although a precise theoretical definition can be elusive (Gambetta 1988, Uslaner 2002, O'Neill 2003, Vallier 2021). Without entering into the theoretical debate on the matter, it is nonetheless worthwhile to point out several methodological problems that arise from a systematic study of trust. As these problems in turn reverberate through questions of measurement, hypotheses of links of causation, and policy prescriptions, it is worth citing four major issues at the outset.

First, trust is ultimately a decision, but it is reached at the end of a process of internal deliberation. There may be tension between the models we ascribe to such deliberation and the bare fact of the decision, which may complicate predictions of trust ("is trust rational?", "are someone's trusted beliefs consistent with one another?", etc.). Second, trust is a belief, which may lead to behavioral effects, but also may not. As we typically observe behavior and infer belief, we are seldom in a position to disentangle cases where what is lacking is the belief, from cases where a belief is latently present, producing no behavior (especially in situations where actors may be less than forthcoming in reporting their beliefs). Third, there is no reason to think that a judgment as to trustworthiness is always front-and-center in actors' minds. Such a state of affairs, however, implies that we may struggle to distinguish the presence of trust from the salience of trust in a specific situation ("is something trusted on its merits, or is the issue of trust never even called into question?"). How salient trust is in different social situations is also context-bound, and thus can depend on cultural and historical factors. Fourth and finally, in aggregate cases it can hardly be guaranteed that the same causal interpretation fit all participants: thus, the efficacy of policy prescriptions hinging on an analysis of trust may be diminished by the multiple different patterns of trust present.

If trust as a freestanding concept can begin to appear evanescent, things become more straightforward when we think in terms of changes in trust. Indeed, many of the real-world cases in which trust is discussed or invoked as an explanatory factor tend to focus on changes in trust. Empirically, much of the data collected on trust is intended to track changes or make comparisons: across institutions, across time, across countries,[2] etc.

The most common procedures through which change is brought about, i.e. social trust is built, or, vice versa, is squandered, are likewise widely discussed in policy circles and public discourse. For instance, trust may be increased by broadening or diversifying the scope of arguments in favor of trustworthiness: verification procedures may be implemented, transparency may be boosted, or perhaps even participation or access on the part of those whose trust is sought may be arranged. Alternatively, different forms of assurance may be given as to future intentions: a systematically earned track record of impartiality in non-trivial cases, or examples of accountability, such as the sacrifice of responsible parties, all the way to major organizational or identity overhauls when the issue appears pervasive and entrenched. Instances of the opposite dynamic are likewise easy to conjure, from steady erosion of a

---

[2]See the chapter on measurement by Bauer & Freitag in Uslaner (2018), as well as the earlier review by Levi & Stoker (2000).

reputation for trustworthiness through a series of marginal missteps to full-blown crises, scandals in which the cognitive dissonance between revelations and reputations has catastrophic consequences.

2 Trust in Contemporary Societies

No discussion of social trust can evade the fact that, in most advanced industrial democracies today, we are living in a period of exceptionally low trust. Such empirical observations cover not only the State and its organs (the government, parliament, the bureaucracy, the courts, the police, the army…) but also the institutions of civil society (political parties, unions, business, the media, churches…); they are robust to different forms of measurement, broadly congruent across countries (in direction if not in scope), and deeply entrenched, so that brief positive shocks soon give way to a regression to low-trust equilibria.[3] The current period, across a range of indicators, is one characterized by the lowest levels of public trust since we have comparable records.

It is important to emphasize that this crisis of trust is not the same phenomenon as political and ideological polarization, although there are correlations in certain national cases (Citrin & Stoker 2018). The crisis of trust is broader, encompassing many more aspects of social interaction that specifically political ones. Its main cleavage is macro-micro: it pits the world of small-scale, face-to-face interaction against larger, impersonal social entities; while personal connections, day-to-day networks, and community-level interactions are less affected, the reputation for trust of institutions in the wider society plummets. Media, inasmuch as it is composed of large-scale organizations, is one of the targeted institutions.

This crisis of trust is especially troubling, since trust is a functional requirement for the operation of complex societies. In particular, given the enormous diversity of our technical endeavors, it is completely unfeasible for individuals independently to validate the systems they use in their daily lives (Giglioli 2020): all no-trust systems rely on trust at some other level (De Filippi et al. 2020), if only as a question of scale. Similarly, thoroughgoing philosophical agnosticism is not compatible with active participation in society: it is necessary to operate by default on the basis of a meaningful distinction between truth and falsehood on objective matters of fact.

It is worth noting that these pragmatic requirements are at odds with hegemonic political and epistemological theories, which insist, respectively, on the right to/advantageousness of disagreement (and thus, also, error), and on the socially constructed, provisional nature of our understanding of the world. Both political liberalism and the modern scientific method are premised on the use and productive regulation of mistrust (Shklar 1998, Rosanvallon 2006, Krastev 2013, Kuhn 1970). Disinformation specifically exploits this mismatch between the ideal open-ended nature of our social knowledge and the actually-existing arrangements prevailing in institutional contexts. For instance, under current conditions science cannot but be experienced by the layperson as embedded within research and industrial institutions of large scale, oriented in whole or in part by capitalist, profit-seeking motives. Its legitimacy is consequently affected.

Trust, like legitimacy, is ultimately not optional at the system level. Therefore, a fundamental tension is potentially able to develop between the pragmatic need to maintain behaviors consistent with default social trust and a more or less widespread belief among the public in the lack of sufficient grounds for such trust. The consequences of this tension between sanctioned behavior and doubt can be witnessed at many different levels. For instance, this state of affairs may be perceived as a

---

[3]The Edelman Trust Barometer has offered a focused time-series of snapshots of this dynamic, but the same results are captured by generalist recurrent polling projects such as Eurobarometer or the Pew Research Center surveys. Also see the Coeuraj report by Parker (2021).

disempowerment of the public, which may react by seeking performative expressions of lack of trust (Munyaka et al. 2022): exhibited mistrust then becomes an identity marker to signal belonging in a perceived out-group. In turn, a society which must contend with a permanently disgruntled cognitive minority necessarily has less degrees of freedom to adapt and govern itself. Maintaining such equilibrium presents considerable risks, especially in crisis situations.

There have been several attempts to account for the contemporary trust deficit.

Structural explanations stress the failure of socialization in interpreting the public sphere, itself induced by broad economic change (globalization) disaggregating many people's social relations. Stated this way, the current crisis is merely an update of the classic social atomization diagnosis (Putnam 2000, Wuthnow 2010). As such, it suffers from all the well-known shortcomings in terms of potential policy solutions. In particular, we know that processes of trust-building are bottom-up, long-term, and based on dense, repeated face-to-face interaction. But the disaggregating forces at work in today's world are precisely top-down, sudden and conjuncture-dependent (in other words, crisis-driven [Konings 2018]), and impersonal. Compared to them, any trust-building exercise will be a remedial, second-order phenomenon.

Alternative explanations resort to technological or cultural arguments. Technological explanations focus on the disintermediation wrought by Information Technology, which has made knowledge once monopolized by professional elites widely available, hence undermining the deference traditionally enjoyed by expertise (Woolley 2023). Cultural explanations, on the other hand, privilege the change in prevailing attitudes in the public, especially in terms of ascribing motivations: a crisis in civic sentiment and a rise in polarization leads to the discounting of any supposed motivation that does not stem from ostensible self- or group-interest, branding any belief in public-mindedness a shameful naïveté (Bronner 2015, Bright 2018, Kavanagh & Rich 2018). Such a line of reasoning is also bolstered by the clear perception that the main terrain of much of economic competition in today's world is asymmetric information, and that therefore public communication can be expected by default to be oriented by the profit motive: the public sphere becomes completely transactional when the pursuit of these informational asymmetries is a core operational strategy.

Whatever the comparative merits of these theories, it is indisputable that the current information ecosystem exists atop a state of crisis in social trust compared to previous generations.

Although the challenge of trustworthiness is felt by institutions everywhere, the means with which the trust deficit is tackled vary significantly by political regime type. It is important to recall for instance that in non-liberal-democratic contexts like the People's Republic of China, granular, public surveillance is openly justified as a means of policing trust, especially in regard to financial matters (e.g., citizenship scores for credit trustworthiness) [Lei 2018, Pei 2024]. By contrast, liberal-democratic societies cannot operate in such a heavyhanded fashion, for their public-policy interventions in bolstering social trust need to balance with the value of privacy.

The democratic public sphere today is vulnerable because it can be 'knocked out of true' in either of two opposite directions: its members can fall into an apathetic disregard for the truth value of any content in the public sphere, or alternatively they can commit passionately, and beyond the possibility of being persuaded otherwise, to the truth of claims incompatible with it. The ideal of the informed democratic citizen exists in this inherently tenuous balance between a desire to verify the truth value of public statements and a willingness to update beliefs as the balance of public evidence shifts. In fact, the traditional information ecosystem is a classic example of an individualist model whose actual functioning in practice was historically guaranteed by the unexamined presence of communal institutions: trusted agents that, though competing with each other, also coordinated mass beliefs whilst (at least in the majority) remaining loyal to the overarching values of the polity. Once this institutional firewall is breached, through social atomization, technology-driven disintermediation, or shifts in

political preferences and media consumption, attacks against individual beliefs gather strength (Risse 2023: chapter 4).

3 Disinformation and Trust

Disinformation today takes place against the backdrop of the trust deficit. Before we begin to address the question of the nature and direction of causal links between trust and disinformation, however, it is useful to examine the latter more closely. The conceptual and measurement difficulties inherent in disinformation are significant, and the debates surrounding them are discussed in other parts of this volume. For present purposes, it is sufficient to mention a subset of these issues, concerning three main points: the relation with pluralism, the importance of intentionality, and material manifestation.

First and foremost, disinformation must be disambiguated from ideological or interest-based disagreement. Given the polarized nature of public dialogue, such a distinction may be fraught. In fact, it is common to expect non-neutral speakers (politicians, corporations, lawyers, PR firms, and so forth) to 'put their best foot forward', stress their side of the case, emphasize the facts that agree with their brief, etc. In a way, their backers and supporters would be satisfied with nothing less. So when does this admittedly rhetorical and self-serving speech devolve into disinformation? When is 'caveat emptor' not warning enough against it? In other words, can the labeling of a certain instance of speech as disinformation depend on the surrounding condition of the media field (Muirhead & Rosenblum 2019)? Specifically, political disagreement cannot simply be reduced to disinformation in liberal-democratic societies. For this reason, purely individual explanatory and analogical models of disinformation, such as individual addiction (radicalization, 'falling down the rabbit hole'…), fail to convince, as they abstract from the ambient level of ideological polarization (Freiling et al. 2023). Such a determination continues to be the case even if public debate does not take place in ideal communicative conditions. The fact, for instance, that organized actors endeavor to influence public opinion by exploiting their influence does not entail that the political or normative position they favor is necessarily nefarious.[4]

The specific difference implied in the notion of disinformation, rather, is an evaluation of the truth contents of information conjoined with (or, in light of) the intention to deceive (Wanless & Pamment 2019). Disinformation has a negative connotation because it does not bespeak participation in an earnest competition between irreconcilable versions of reality by their respective institutional backers: there is always something underhanded, something surreptitious about disinformation. Disinformation is always asymmetrical ideological struggle. For this reason, it is typically couched as a disagreement of fact, while in actuality its animus is driven by disagreements of value.

The purpose of disinformation is clear, but its material form may be hard to isolate or define. Such a difficulty follows from disinformation being a hack, an exploit: the mechanics of application in the individual case do not ultimately matter, or define its essence, as much as its rate of success.[5] Inasmuch as disinformation is an active and purposive strategy, the goal of its perpetrators must be not simply to change beliefs, but also (immediately or in some future moment) to change behavior in consequence. Hence, while we can imagine (as we have mentioned above) mistrust as a latent phenomenon, a background attitude without a direct manifestation, the point of disinformation is to be performative.

---

[4]Such a behavior is, however, necessarily a straining of the rules of the game of public discourse, which requires openness and transparency in the authorship of public interventions. In other words, process matters for public discourse's integrity, especially in terms of the trustworthiness of participants.

[5]For this reason, it is imperative that a theoretical approach be able to distinguish factors that affect the production and dissemination of disinformation from factors that facilitate its success.

Similarly, disinformation is a form of predation: in this regard, it is most akin to a confidence scam (Maurer 1999, Nash 1976, Goffman 1952), or to social engineering (Mitnick & Simon 2002, Gehl & Lawson 2022). Therefore, one would expect there to be a difference in types of disinformation based on whether the target is broad or narrow, chosen on the basis of general demographic or of personally identifiable characteristics, etc.

From the standpoint of this brief characterization of disinformation as an aggressive pathology of the information ecosystem, we can turn to the question of whether it represents a symptom or a cause of the contemporary trust deficit. As will be discussed below, some of the major problems that trust poses for disinformation, both conceptually and at the policymaking level, derive from factors that the characterizations presented so far have sought to highlight. While the contemporary crisis of trust is a structural phenomenon, attaching itself to most all social institutions and therefore presenting itself essentially as a *Zeitgeist*, a mentality, shared by a culture as a whole, in disinformation as we have discussed it, i.e. as a speech act, the role of the individual actor is crucial.

4 Patterns of Causation

The idea of a causal relationship between the lack of social trust and disinformation is intuitively appealing, but it can be challenging to corroborate conclusively.[6]

It is certainly possible that the dynamics of causal influence run in both directions. Likewise, there may be correlations in long- or short-term shifts in the two variables, with or without significant lags. Moreover, there may be a symmetry in one direction (e.g., increased mistrust begetting higher prevalence of disinformation) but not in the other (i.e., decreases in the two variables)—which would have notable implications for policy intervention, for it would suggest that, for instance, preventative approaches be preferred to remedial ones.

Although we may think that the most plausible link between mistrust and disinformation is a feedback loop or vicious circle, occurring as a race to the bottom in the context of a delegitimation cycle, it is still preferable analytically to explore the two directions of causation as distinct. Therefore, if we conceive of disinformation as a tactic, we can reach different interpretations of it based on whether it is a consequence or a precondition of low trust. Each of these options gives rise to two working hypotheses.

Specifically,

(a) if it is disinformation that causes low trust:

– Hp 1: such actions are intended directly as an attack on the trustworthiness of the target institution.

Alternatively,

– Hp 2: each individual action is merely one of a myriad of self-serving behaviors that trade on the reputation of the target institution: the cumulative result of falling trust in the target institution is an unintended consequence.

Therefore, if Hp 1 presents the case of disinformation as surgical informational warfare, Hp 2 explores the emergent properties of uncoordinated collective action.

(b) If, on the contrary, disinformation is a consequence of low trust:

– Hp 3: institutions with a trust problem, i.e. a significant trust deficit, tend to attract disinformation attacks.

Alternatively,

---

[6]Some interesting recent studies include: O'Connor & Weatherall 2019, Howard 2020, Tripodi 2022, Berinsky 2023.

– Hp 4: disinformation is a broadly-targeted phenomenon that is endemic everywhere, but its rate of success spikes in cases in which trust is weakened.

The distinction between Hp 3 and Hp 4 depends on the visibility of social trust, which translates into an interpretation of disinformation as an opportunistic behavior vs. a mostly blind or automated process, whose ultimate effectiveness nonetheless has cognitive implications for the actual trustworthiness of institutions.

The purpose of presenting these hypotheses here is not to adjudicate them empirically, nor should we think that the answer may not differ in specific cases. The purpose, rather, is to set up a typology for the interpretation of our current knowledge and for the orientation of further empirical inquiry. This approach, in turn, will force more clarity about the causal models underpinning suggested policy interventions, especially in terms of being able to diagnose the causes of policy failures.

Hence, if the empirical evidence points to the fact that a specific instance falls into case (a) [Hp1 or Hp2], policymaking on disinformation will be targeting the root cause, rather than the epiphenomenon, but will also be invested with significantly weightier responsibilities hanging in the balance. If the case falls under Hp1, the policymaker might want to focus on retaliatory mechanisms against single large disinformation actors, for instance by facilitating the victims' legal ability to recover downstream reputational losses from the perpetrators. If, on the contrary, Hp2 is corroborated by the evidence, the policy implication would be to emphasize broader structural measures, to dampen the temperature of debate or impose costs on participants to 'thin out' the disinformation ecosystem.

Vice-versa, in the other branch of the tree, case (b) [Hp 3 or Hp4], policymaking on disinformation will be an indirect affair, because modifying social trust is often not a lever directly available for policymaking manipulation, at least in the short run. Nonetheless, determining whether Hp3 or Hp4 holds would suggest a response strategy based either on attacking specific institutional pathologies (Hp3), hence structural rearrangements, changes in leadership, rebranding, etc., or else on more general trust-enhancing initiatives (Hp4), hence focusing on a whole-of-society approach.

In the same vein, it may prove useful to interrogate other aspects of the trust-disinformation link. For instance, as we have discussed above there may be different drivers of mistrust, attaching to different actors, with different motivations; do such differences correlate with different forms of disinformation, or different rates of its success?

To consider a concrete example, the most common accusations against mistrusted institutions are political bias and corporate special interests. Do these two imputed motives make particular forms of disinformation about these institutions more believable?

Similarly, is the class of institution that is mistrusted of any importance in considering the shape of disinformation leveled at it? Is there perchance a particular susceptibility of media organizations (or of tech companies, inasmuch as they take up the role of media gatekeepers)? In this case, the general pattern of media mistrust can be interpreted as a calling into question of the media's speech position as a neutral or disinterested arbiter of the public sphere. The implication then would be that impartial speech acts in public discourse are rendered impossible for all intents and purposes, because the pervasive search for hypocrisy and hidden motives is always looming.

Ultimately, beyond the mere fact of an involvement of trust deficits in the use of disinformation, an analysis of the origin, prevalence, shape, and expression of such deficits in specific instances may prove advantageous empirically in the process of devising a policy response to disinformation.

5 Trust, Disinformation, and Policymaking

Turning from conceptual analysis to a consideration of policy implications, it is necessary to stress at the outset the centrality of institutions. If disinformation is to be stemmed, the most straightforward and

efficient intervention will require the active participation of institutions. Hence, trust in institutions will always be relevant to disinformation and to the information ecosystem. Indeed, asking whether disinformation can be fought is akin to asking whether there exists in society any institution or coalition of institutions credible enough to rebut and debunk it. Conversely, to say that an institution does not possess the material or immaterial resources necessary to defend itself against a disinformation or delegitimation attack is tantamount to saying that it is a weak, even a powerless institution: its trust deficit will simply coincide with reputational disdain.

However, such policy engagement may need to take place within a broader context of low trust. In other words, institutions might need to take their perceived level of trustworthiness as a given at any one point in time, reputational change occurring on a different temporal scale from a disinformation attack, if it even is in the power of the institution to effect in the first place (Lewicki & Brinsfield 2017). Moreover, the ability of any collective social actor to function as a trust builder in today's world is typically an externality: it is seldom a priority for any one organization, and is necessarily trumped by other material incentives. Consequently, the ability to implement policy fixes centered on trust is hampered, and the most relevant policy tools for combating disinformation at present refer to the spectrum of low-trust equilibria.[7]

One distinction that is worth mentioning can be borrowed from criminology: one must move from a mindset of eradication to one of risk management, understood as actuarial minimization. In other words, the policy goal should be to govern a world in which disinformation continues to be present, by preventing major disinformation success. At a minimum, it is possible to define disinformation success in public policy as a pronounced shift of the realm of the politically thinkable on the basis of injection into public debate of relevant information that is purposely false. Therefore, the minimalist policy program for countering disinformation centers on preserving the viability of the public sphere.

The problem, of course, is that disinformation is heavily stigmatized: it amounts to a denunciation of insincerity or lying, and as such is definitive and decisive in our culture of discourse: once an interlocutor has been labeled as a liar and a cheat, the stakes are immediately systemic; there is no further common ground on which to continue a conversation.[8] Either such a denunciation has the power to provoke consistent consequences and ostracize the interlocutor so labeled, or else the accuser proves both weak and unprincipled by seeking a modus vivendi, an accommodation with an opponent branded as transgressing social norms.

The problem is compounded if there is no way to separate the perpetrators of disinformation from major domestic social or political forces. The general perspective we have entertained so far is the presence of threat actors attempting to undermine institutions in the public sphere, who in turn battle to retain their trustworthiness. However, it is not difficult to imagine a contingency in which institutions themselves (or their leaders at any one time) might act as spoilers of trust in the public sphere. In this scenario, any communicative act becomes an attempt at manipulation, for the institution (or its leaders) does not much care for its own reputation (or perhaps no longer suffers significant reputational sanctions for disinformation attempts), and a negative-sum dynamic develops with regard to the believability of any information in the public sphere (Benkler et al. 2018, Cosentino 2020).

Given this possibility, it is clear that the social distribution of mistrust is decisive. Inasmuch as the main social stakeholders can be kept (even artificially) at arm's length from actual disinformation operations, accountability might suffer, but public discourse as a whole stands a chance not to devolve into a zero-sum game.

---

[7]Technology solutions that go in the direction of 'zero-trust' remedies, for instance in the field of cybersecurity or of digital identity, reflect precisely this reality (Young 2020, Preukschat & Reed 2021).

[8]On the socially isolating effffects of lying, see Koyré (2016). Also see Lynch (2004), especially chaps. 9-10.

This primacy of the preservation of the public sphere and its viability is a specifically liberal-democratic goal, because, as has been mentioned, in fighting disinformation these polities also care about preserving other values and norms[9]. In turn, this respect for values, even as aspirational regulative ideals, is one of the most profound, but also effective, claims to institutional and public trust for liberal regimes (Waldman 2018). For authoritarian regimes, on the contrary, the fight against disinformation is a purely technical matter: these regimes resign themselves from the start to a low-trust public environment, with all its informational consequences (the omnipresence of rumors, the plague of self-serving reporting up the chain of command and consequent distortion of the elite's understanding of objective conditions, and so forth) [Figes 2007, Glaeser 2011, Pomerantsev 2014].

To conclude, what are the practical implications of the foregoing analysis, given that the secular drivers of mistrust in democratic societies do not seem likely to be curbed in the near term? It is possible to sketch two possible scenarios. In the first, we witness a stable low-trust equilibrium, in which an erosion of the public sphere as we know it entrenches polarization but does not precipitate an implosion of economic activity or market forces, governance manages to salvage basic levels of service delivery, and civil life is not violently impacted. In the second, worst-case scenario, a catastrophic breakdown in social trust results in a dramatic crisis in basic social processes, with widespread spillover into real-world consequences.

In either case, considering the problem of disinformation from the perspective of social trust yields sobering insights about the fragility of our polities, the need for careful and pragmatic management, and the imperative to responsibility that falls upon all actors in the information ecosystem.

---

[9]In fact, Farrell & Schneier (2018) suggest that autocracies and democracies, conceived as informational systems, have different and distinct attack surfaces and threat models.

References

Banfield, E. C., & Banfield, L. F. (1958). *The moral basis of a backward society.* The Free Press.

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Berinsky, A. J. (2023). *Political rumors: Why we accept misinformation and how to fight it*. Princeton University Press.

Bright, J. (2018). Explaining the Emergence of Political Fragmentation on Social Media: The Role of Ideology and Extremism. *Journal of Computer-Mediated Communication, 23*(1), 17–33. https://doi.org/10.1093/jcmc/zmx002

Bronner, G. (2015). *Belief and misbelief asymmetry on the internet*. ISTE Ltd/John Wiley and Sons Inc.

Citrin, J., & Stoker, L. (2018). Political Trust in a Cynical Age. *Annual Review of Political Science*, 21, 49-70.

Cosentino, G. (2020). *Social media and the post-truth world order: The global dynamics of disinformation*. Palgrave Macmillan - Springer Nature.

De Filippi, P., Mannan, M., & Reijers, W. (2020). Blockchain as a confidence machine: The problem of trust & challenges of governance. *Technology in Society, 62*. https://doi.org/10.1016/j.techsoc.2020.101284

Farrell, H., & Schneier, B. (2018). Common-Knowledge Attacks on Democracy. *Berkman-Klein Center Research Publication,* No. 2018-7. https://ssrn.com/abstract=3273111

Figes, O. (2007). *The whisperers: Private life in Stalin's Russia*. Allen Lane.

Freiling, I., Krause, N. M., & Scheufele, D. A. (2023). Science and Ethics of "Curing" Misinformation. *AMA Journal of Ethics*, *25*(3), E228-237. https://doi.org/10.1001/amajethics.2023.228

Fukuyama, F. (1996). *Trust: The social virtues and the creation of prosperity* (1. Free Press paperback ed). Free Press.

Gambetta, D. (Ed.). (1988). *Trust: Making and breaking cooperative relations*. B. Blackwell.

Gehl, R. W., & Lawson, S. T. (2022). *Social engineering: How crowdmasters, phreaks, hackers, and trolls created a new form of manipulative communication*. The MIT Press.

Giglioli, M. (2020). *I labirinti della sorveglianza informatica: Cittadinanza e impegno politico nell'era della trasparenza universale*. Il mulino.

Glaeser, A. (2011). *Political epistemics: The secret police, the opposition, and the end of East German socialism*. The University of Chicago Press.

Goffman, E. (1952). On cooling the mark out: Some aspects of adaptation and failure. *Psychiatry: Journal of the Study of Interpersonal Relations, 15*(4), 451–463.

Howard, P. N. (2020). *Lie machines: How to save democracy from troll armies, deceitful robots, junk news operations, and political operatives*. Yale University Press.

Kavanagh, J., & Rich, M. D. (2018). *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life*. RAND.

Konings, M. (2018). *Capital and time: For a new critique of neoliberal reason*. Stanford University Press.

Koyré, A. (2016). *Réflexions sur le mensonge*. Allia.

Krastev, I. (2013). *In mistrust we trust: Can democracy survive when we don't trust our leaders?* TED Conferences. https://www.overdrive.com/search?q=88577B13-83F7-4362-B961-D1F7067CEA8D

Krueger, F., & Meyer-Lindenberg, A. (2019). Toward a Model of Interpersonal Trust Drawn from Neuroscience, Psychology, and Economics. *Trends in Neurosciences, 42*(2), 92–101. https://doi.org/10.1016/j.tins.2018.10.004

Kuhn, T. S. (1970). *The structure of scientific revolutions*. University of Chicago Press.

Lei, Y.-W. (2018). *The Contentious Public Sphere: Law, Media, and Authoritarian Rule in China*. Princeton University Press.

Levi, M., & Stoker, L. (2000). Political Trust and Trustworthiness. *Annual Review of Political Science, 3*, 475–507.

Lewicki, R. J., & Brinsfield, C. (2017). Trust Repair. *Annual Review of Organizational Psychology and Organizational Behavior, 4*, 287-313.

Lynch, M. P. (2004). *True to Life: Why Truth Matters*. MIT Press.

Mann, M. (2012). *The sources of social power. Volume 1, A history of power from the beginning to AD 1760* (New ed). Cambridge University Press. https://doi.org/10.1017/CBO9781139381307

Maurer, D. W. (1999). *The big con: The story of the confidence man*. Anchor Books.

Mitnick, K. D., & Simon, W. L. (2002). *The art of deception: Controlling the human element of security*. Wiley.

Muirhead, R., & Rosenblum, N. L. (2019). *A lot of people are saying: The new conspiracism and the assault on democracy*. Princeton University Press.

Munyaka, I., Hargittai, E., & Redmiles, E. (2022). Misinformation Paradox: Older Adults are Cynical about News Media, but Engage with It Anyway. *Journal of Online Trust and Safety, 1*(4). https://doi.org/10.54501/jots.v1i4.62

Nash, J. R. (1976). *Hustlers and con men: An anecdotal history of the confidence man and his games*. M. Evans.

O'Connor, C., & Weatherall, J. O. (2019). *The misinformation age: How false beliefs spread*. Yale University Press.

O'Neill, O. (2003). *A question of trust*. Cambridge University Press.

Parker, N. (2021). *Navigating Trust in US Society*. Coeuraj USA.

Pei, M. (2024). *The sentinel state: Surveillance and the survival of dictatorship in China*. Harvard University Press.

Pomerantsev, P. (2014). *Nothing is true and everything is possible: The surreal heart of the new Russia*. PublicAffairs.

Preukschat, A., & Reed, D. (2021). *Self-sovereign identity: Decentralized digital identity and verifiable credentials*. Manning.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.

Putnam, R. D., Leonardi, R., & Nanetti, R. Y. (1992). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press.

Risse, M. (2023). *Political theory of the digital age: Where artificial intelligence might take us*. Cambridge University Press.

Rosanvallon, P. (2006). *La contre-démocratie: La politique à l'âge de la défiance*. Seuil.

Shklar, J. (1998). The Liberalism of Fear. In S. Hoffmann (ed.), *Redeeming American Political Thought*. University Of Chicago Press.

Tilly, Charles. (2005). *Trust and rule*. Cambridge University Press.

Tripodi, F. B. (2022). *The Propagandists' Playbook: How conservative elites manipulate search and threaten democracy*. Yale University Press.

Uslaner, E. M. (2002). *The moral foundations of trust*. Cambridge University Press.

Uslaner, E. M. (Ed.). (2018). *The Oxford handbook of social and political trust*. Oxford University Press.

Vallier, K. (2021). *Trust in a polarized age*. Oxford University Press.

Waldman, A. E. (2018). *Privacy as trust: Information privacy for an information age*. Cambridge University Press.

Wanless, A., & Pamment, J. (2019). How Do You Define a Problem Like Influence? *Journal of Information Warfare*, *18*(3).

Woolley, S. C. (2023). *Manufacturing consensus: Understanding propaganda in the era of automation and anonymity.* Yale University Press.

Wuthnow, R. (2010). *Be very afraid: The cultural response to terror, pandemics, environmental devastation, nuclear annihilation, and other threats*. Oxford University Press.

Young, K. (2020). *The domains of identity.* Anthem Press.