THE
CONSILIENCE
PROJECT

# The Endgames of Bad Faith Communication

FOUNDATION    **23 Feb 2022**  |  9 min read



*Seeking to understand others and communicate honestly is an essential democratic virtue. Can it be maintained in the digital age?*

Decades of culture war have degraded civic discourse, putting many open societies in a tailspin of bad faith public communications. Politicians, journalists, and everyday people on all sides intentionally mislead with facts, mischaracterize opposing views, and dehumanize those with whom they disagree.[1] Social media has started to change our basic habits of communication by amplifying and incentivizing bad faith tactics. Every day, whole populations are exposed to powerful forms of computational propaganda and other manipulative information.[2]

*Bad faith communication has become normalized.* This is bad news for any society that values and seeks to rely upon the uncoerced cooperation of its members. History tells us that the endgames of society-wide communication breakdowns are catastrophic. When open communication cannot be used to resolve conflict and coordinate behavior, societies are driven towards chaos, war, oppression, and authoritarianism. Restoring public trust in good faith communications is possible. But it requires both a cultural shift toward civic virtues and a redesign of the technologies and social processes that structure civic discourse.

## People are being deskilled in the art of good faith communication, while refining skills in bad faith tactics.

Despite the new normal of widespread bad faith communication, good faith communication is

towards shifting the cultural balance back in its favor.[3]

Communicating in good faith as a society requires that people take up certain skills and commit to shared values. These skills and values are generally not practiced and endorsed in most contexts of civic discourse today. People are instead being deskilled in the art of good faith communication, while refining skills in bad faith tactics (see both Box 1 and 2 below).

**Good Faith Communication:** discourse oriented towards mutual understanding and coordinated action, with the result of increasing the faith that participants have in the value of communicating.

**Some Signs of Good Faith Communication:**
- Expressions of humility and curiosity
- Openings for changes in position based on new information
- Disagreements welcomed; group learning valued
- Steelmanning the position of others
- Respect maintained during disagreement
- Sufficient time given to open discussion and other aspects of fair process
- Use of careful clarifications and evidence
- Attempts at finding shared base realities and values
- Emergence of new positions, integrations, and nuance

*Note: All signs of good faith communication can be "faked" in bad faith.*

**Bad Faith Communication:** discourse that is intended to achieve behavioral outcomes (including consensus, agreement, "likes") irrespective of achieving true mutual understanding, with the result of decreasing the faith participants have in the value of communicating.

**Some Signs of Bad Faith Communication:**
- Expressions of hubris and lack of curiosity in opposing views
- Refusing changes in position based on new information
- Disagreements unwelcomed; consensus overstated
- Strawmanning the position of others
- Disrespect included as part of disagreement
- Insufficient time and other aspects of unfair process
- Avoidance or omission of careful clarifications and evidence
- No attempts to find shared base realities and values
- Emergence of stalemates, polarization, and simplifications

*Note: All signs of bad faith communication can be disguised and denied.*

The situation has degraded to the point where it is widely believed that calls to good faith (such as this paper) are themselves acts of bad faith, undertaken only by those interested in controlling the discourse. Calls for good faith communication are understood at best as naive requests to calm the outrage and conflict that now runs rife in political discourse. Both ends of the political spectrum (the far left and the far right) express this view. Both sides believe that "the other side" simply can't be trusted and therefore cannot be engaged in good faith. To do so would be to fall into a trap, serving only to validate the dangerous views of groups known to be acting in bad faith.

This stance of assuming the undesirability (and sometimes impossibility) of good faith communication sets off a spiral of mutual dismissal, distrust, and villainization. The only outcome of this dynamic is escalating cultural conflict and, eventually, physical violence. A growing awareness of this dynamic has political scientists and journalists worrying about increasing violent civil conflict within the United States. Others see the situation leading to some form of authoritarian control over public communications to secure social coordination by force.[4]

## There should be no illusion: today's culture war cannot be won by any side.

All sides agree that in the absence of good faith communication the last resort is some form of violence. Despite this knowledge, many still assert that it is actually unethical to engage "the other side" in good faith. The result is that people are earnestly doing something they believe to be right—refusing to engage in good faith with those they disagree with—which is nevertheless leading towards a result they do not actually want. Meanwhile some parties actively seek to benefit from this dynamic, such as social media companies that leverage conflict for attention capture.[5] The culture war, like most wars, is a source of profit, and is therefore perpetuated despite the dangers.

Given well-documented advances in the field of information warfare, there should be no illusion: *today's culture war cannot be won by any side*. Mutually assured destruction is now the name of the wargame.[6] The saturation of bad faith communication throughout culture is steadily increasing, like a kind of dangerous background radiation emitted from scientifically engineered memetic weaponry. Public political discourse is quickly becoming a toxic warzone, leaching externalities into families, friendships, and identity structures.

## When institutions that depend on public trust indulge in bad faith communications repeatedly, the legitimacy of these institutions declines.

While insincere and manipulative communications have always been a part of human societies, our modern communications environment surpasses most recent historical periods in terms of the scale of its dysfunction. Comparisons may only be drawn with epochs characterized by failed states, civilizational collapse, and total war. Widespread inability to reach mutual understanding between members of the same society leads inevitably to social breakdown. When institutions that depend on public trust indulge in bad faith communications repeatedly, the legitimacy of

# THE CONSILIENCE PROJECT

## Common Strategies of Bad Faith Communication

| | |
|---|---|
| **Misleading with facts** | Presenting an argument containing factual information, which is used intentionally to lead others to draw a conclusion that is not entirely accurate. |
| **White hat bias** | Presuming one's own moral and intellectual correctness, then using that assumption as righteous justification for communications that are intentionally deceptive and manipulative. |
| **Strawman arguments** | Presenting the arguments of opponents in their weakest forms, and after dismissing those, claiming to have discredited their position as a whole. |
| **Ad hominem dismissal** | Disparaging the character or person of others, and in so doing acting as if this also invalidated their arguments. |
| **Moving the goalposts** | Establishing an agreed standard (criteria or data) for accepting others' views, but once this is provided or met, the prior agreement is not mentioned, and a new standard is set. [The reverse case also applies, i.e., when one cannot meet the agreed standards, these standards are forgotten, and new ones are established.] |
| **Sanctimony** | Acting as if oneself and/or group is unquestionably morally superior and more intelligent than specific disagreeable individuals or groups, and thereby devaluing the members and delegitimating all the views of that group. |
| **Appeals to authority** | Deeming that one's own authority, that of a favorite expert, or that of an associated institution, definitively establishes positions currently being contested, therefore no further communication or explanation is needed and the arguments of the disagreeable parties can be dismissed. |
| **Dehumanizing language** | Deploying language that characterizes groups as irredeemably unreasonable and not worthy of consideration, and thereby suggesting such groups should not be engaged in good faith. |

| | |
|---|---|
| **pressure** | conformity override the power of reason and evidence. This includes "canceling," deplatforming, unfollowing, blocking, boycotting, trolling, etc. |
| **Pejorative representations** | Employing openly insulting and dismissive language when describing the persons, ideas, or practices of disagreeable groups, thereby justifying the discounting of their arguments without earnest consideration. |
| **Faking empathy and respect** | Pretending to feel empathy and respect for disagreeable others in a manner that undermines their actual experiences and beliefs— "strawman empathy." |
| **Equivocations and false logics** | Engaging involved and detailed forms of argument that are nevertheless fallacious and misleading due to subtle (and not so subtle) logical mistakes, such as strategically conflating and misusing terms (equivocation). |
| **Manipulative framing** | Using metaphors and emotional frames to lead preemptively to conclusions that are not fully suggested by the details of the argument. |
| **Villainization** | Creating the image of an "anti-hero" who epitomizes the worst of the disagreeable group, and contrasts with the best qualities of one's own, then characterizing all members of the other group as if they were identical to that image. |
| **Oversimplification** | Intentionally focusing on only a few (or the wrong) variables when drawing conclusions about complex systems, while also dismissing as irrelevant or misleading the views of those seeking to include more variables for consideration. |
| **Complexity smoke screen** | Bringing an overwhelming amount of complex information to an argument and in so doing strategically downplaying a smaller, less complex set of variables that are actually more meaningful to the topic under discussion. |

**Box 2**

In the wake of digital communication technologies—especially social media—manipulative
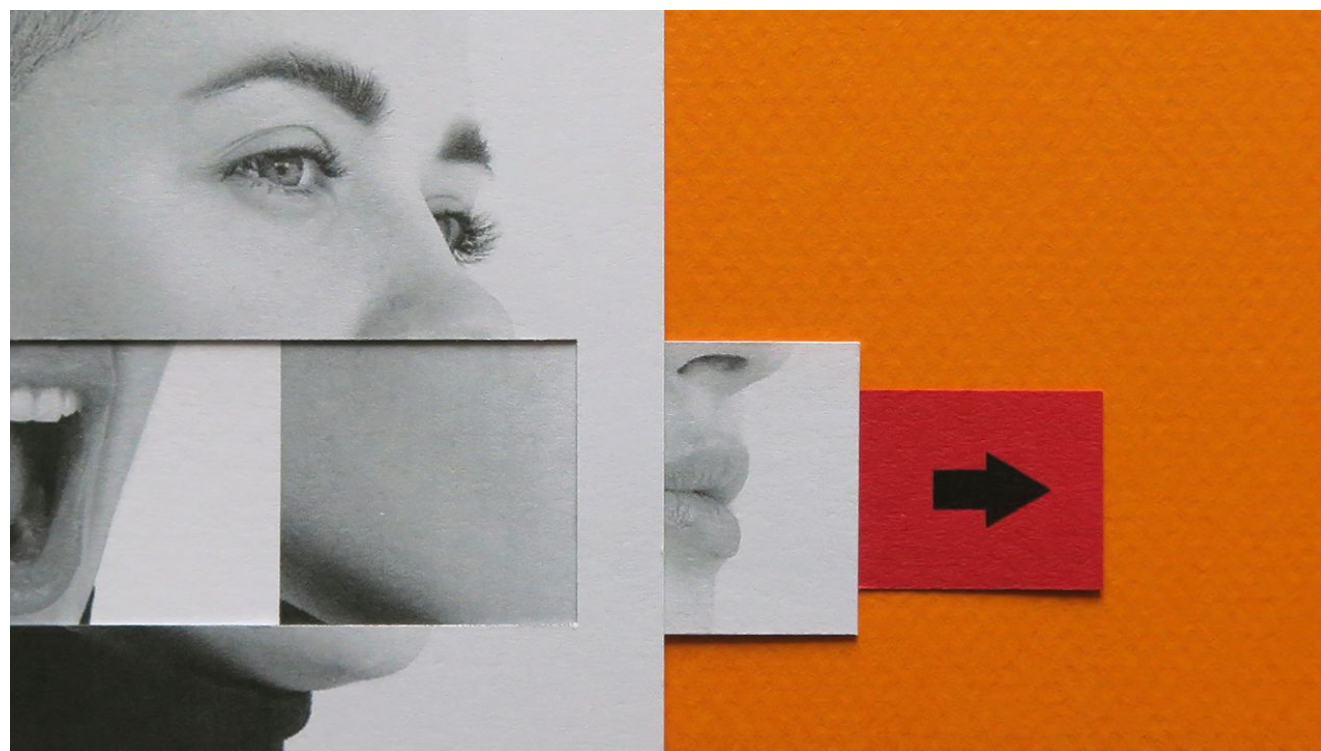
# THE CONSILIENCE PROJECT

## A "post-truth" culture is a culture of bad faith.

A key feature of escalating extremism is a belief that group membership requires bad faith engagements with out-groups. In these contexts, bad faith behavior is often justified to maintain in-group membership and consensus. The normalization of bad faith communication contributes to the creation of extreme in-group pressures, which can rupture identities and exacerbate mental health crises. Personal instabilities usually lead to a doubling down on the need for group membership, increasing rationalizations and amplifications of bad faith practices.[8]

Digital media companies' business models result in a proliferation of increasingly niche group memberships. They also incentivize public displays of conflict and bad faith communication, in order to capture attention and optimize engagement. Advertisements and propaganda dominate the social media space, driving up the total amount of bad faith communication to which people are exposed.[9] In a very literal sense, heavy users of social media are being behaviorally entrained to engage disproportionately in bad faith communication. Politicians, public officials, and influencers of all kinds seek to exploit this environment of distrust and capitalize on the declining social value of good faith interactions. The epistemic commons is repeatedly degraded to the point of exhaustion.[10] A "post-truth" culture is a culture of bad faith.

This must stop if there is to be any future for open societies. Although there are many significant barriers to acting in good faith, they are surmountable given sufficient interest and willingness to seek cooperation and mutual understanding. If the desired outcome is sustainable uncoerced social cooperation—which is what democracies strive to achieve—then willingness and interest must be found.

## The remedy for ongoing bad faith communication is not more bad faith communication.

# THE CONSILIENCE PROJECT

emotional stance leading to an overzealous search for agreement. In part, this is because naive good faith interactions can often be based on an innocent orientation towards peacemaking and the avoidance of conflict. This can be seen in the good faith engagement of a child who does not know that some people are untrustworthy. Unskillfully engaging in good faith can be a danger when others are acting in bad faith.

In some political discussions today, it is common to hear that "you can't engage in good faith with Nazis!" A generous interpretation of this statement is that there are truly unreasonable people who must not be trusted because they have proven themselves to be dangerous and unethical. But this is not an argument against good faith communication in general. It's an argument against engaging in naive forms of good faith communication, which would play into the hands of those actively seeking to cause harm to others. Likewise, for those who have been lied to historically and treated with disrespect by specific groups, naively engaging with those same groups again in good faith would be foolish.

Still, it remains true that under the conditions of contemporary society the endgame of unrelenting and escalating bad faith communications is one in which everyone loses. *The remedy for ongoing bad faith communication is not more bad faith communication*. The risk of mutually assured destruction in the information wargame cannot be contained. The "solution" therefore appears to be necessarily drastic and unavoidably painful. Without a change in the current trends, authorities are increasingly likely to use force to secure ongoing social coordination. This includes increasingly overt censorship and the ideologically motivated disenfranchisement of dissidents.

Avoiding social catastrophe will require the deployment of *highly skilled, non-naive good faith communication*. This is a form of communication that seeks to increase for all parties their faith in the value of ongoing communication. Communicating in this way involves recognizing conflict and disagreement as real and important, unlike less skilled forms of good faith communication, in which conflict is judged to be bad and is defused. The goal is to avoid the downward spiral of mutually reinforcing escalations of bad faith. Delicately transforming a situation of escalating bad faith requires the slow establishment of previously unrecognized shared interests, often on issues as basic as self-preservation. The goal in most cases is not agreement—that would be naive—the goal is simply to preserve the possibility of communication itself. More specifically, the goal is to act in ways that generate faith in the value of ongoing future communication. This is faith in a dynamic of communication that makes it possible to change positions, learn, and improve mutual understanding about essential shared realities.

The risks of engaging in good faith can be mitigated. You can strongly disagree with someone in good faith. But this can only happen if you are in the right context, while also being skillful and committed enough to do so. The only way to handle highly charged disagreements in good faith is to have social processes robust enough to hold them. Social media is demonstrably not adequate to this task. But neither are the legacy institutions—such as the judiciary, legislature, and election process—built by the architects of the modern nation state. Digital communications technologies make possible new forms of democracy, as well as new civic communications infrastructures supportive of the good faith exchange essential to open societies. But these possibilities remain unactualized.

Designing technologies that enable good faith communication is not enough; also required is a culture that emphasizes its value as a shared social good. Once underway, good faith communication creates a self-reinforcing dynamic of mutual understanding—*even during disagreements*—that keeps open possibilities for learning and coordinated action. This is the reverse of the downward spiral of mutual distrust created by bad faith communication. Quick returns of social benefit result from investments in good faith communication across disagreements.
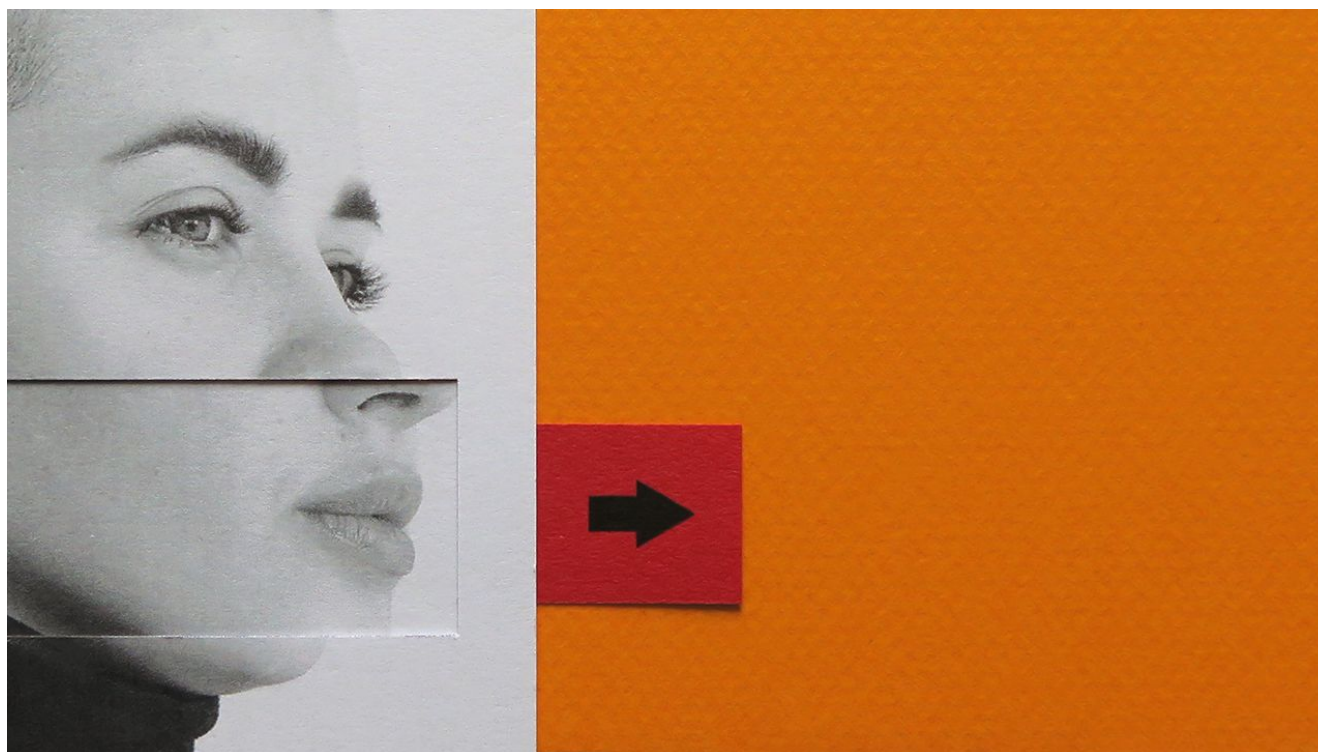


© Adaptive Cultures

Finally, even though most contexts do not currently support it, there is some personal responsibility that can be taken. Despite the escalating culture war, individuals can cultivate the skills of good faith communication and demonstrate its value as a social good. You can practice disagreeing in ways that maintain relational integrity and respect. Reflect on values. Notice bad faith tactics. Hold politicians, officials, journalists, your friends, and yourself to a higher standard of communication.

To get started you could consider the common strategies of bad faith communication in Box 2 and practice the opposing strategies for good faith. But understand that highly skilled, non-naive good faith communication cannot be made routine. There is not a formula or practice or technology that assures it—and nor should there be. Any approach that becomes a recognized signal of "good faith!" will then be faked in bad faith. Individuals must therefore continually innovate in their approach to communication. We must work together always to find new ways to break the hegemony of bad faith. This should be done as if the future of civilization depends on it—because it does.

# Footnotes

1.   See Consilience Papers, "How to Mislead with Facts," *The Consilience Project*, January 30, 2022. ↩

4.   See Consilience Papers, "Teetering Between Oppression and Chaos," July 18, 2021. ↩

5.   See Consilience Papers, "Social Media Enables Undue Influence," December 5, 2021. ↩

6.   See Consilience Papers, "It's a MAD Information War," July 25, 2021. ↩

7.   See Consilience Papers, "Social Media Enables Undue Influence," December 5, 2021.. ↩

8.   For more on what happens when individuals in extremist groups meet with disconfirming evidence, see: Joel Dimsdale, *Dark Persuasion: A History of Brainwashing from Pavlov to Social Media* (Yale University Press, 2021). ↩

9.   See Consilience Papers, "It's a MAD Information War," July 25, 2021. ↩

10.   See Consilience Papers, "Democracy and the Epistemic Commons," February 27, 2021. ↩

## Subscribe to our newsletter →

The Consilience Project is a publication of the Civilization Research Institute (CRI), a charitable think tank focused on reducing systemic fragility and advancing new forms of governance and coordination. CRI seeks to inform the next era of human collaboration and collective intelligence to ensure human and ecological thriving.

A CRI Initiative

**About**
**Team**

If you want to help, have a special skill, are part of a partner organization, or want to help spread the word, please reach out. We value your feedback and support.