



# Toxic Online Environments are what Makes Rational Persuasion Become Wrongful

Lavinia Marin<sup>1</sup>

Received: 22 March 2024 / Accepted: 24 March 2024 / Published online: 27 March 2024  
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Suppose you wanted to switch to a healthier diet and searched for ideas online. You will unavoidably find some posts and videos where various influencers weigh in on this topic on social media platforms. Going down the rabbit hole of online searches, you see that the most vocal debate occurs between two camps: the high fat versus the high carbohydrates. In these debates, a key argument is what counts as a blood marker for health assessment. Some say it is the LDL cholesterol, while others say that it is the triglycerides/HDL ratio. You are no medical expert (otherwise, you would not be searching for this online), so both technical debates are equally unfamiliar to you. Both sides can cite scientific studies showing that LDL and triglycerides/HDL are important health markers and that one is more relevant. If you were to side with any of the two camps, you would make a rational choice, given that both sides have factual and scientific arguments to support their claims. But depending on which side you are persuaded by, your diet will change radically, which will have long-term effects and possibly disastrous consequences for your health.

In deciding which side you take, the scale is not tipped by the strength of the reasons—since you cannot decide that—but on minor circumstances such as which side gets to present their reasons to you first. If it changes your attitude towards a healthy diet, this presentation of reasons will count as rational persuasion, following the schema provided by Mitchell and Douglas (2024). In their paper titled “Wrongful Rational Persuasion Online” (2024), Mitchell and Douglas offer a strict account of rational persuasion, defined as the successful “influencing another by providing reasons” (p. 35) and based solely on those reasons. Reasons are understood as factual, true claims about the world. It would seem that no harm is done if you change your attitude based on reasons alone, right? Not quite. In several circumstances, rational persuasion, previously considered good or at least morally neutral, can be harmful, as Mitchell and Douglas (2024) argue. Some of the conditions for wrongful rational persuasion include audience tailoring, having the wrong standing to persuade (Should an influencer weigh in on this matter at all?), and persuading someone to

---

✉ Lavinia Marin  
l.marin@tudelft.nl

<sup>1</sup> Ethics and Philosophy of Technology Section, TU Delft, Delft, the Netherlands

believe in a falsehood (by tailoring reasons and truncating technically true information). In their paper, Mitchell and Douglas show that most conditions for wrongful rational persuasion are met, especially by online social environments such as social media platforms.

Their contribution is novel insofar as it goes against the mainstream philosophical arguments that see only persuasion based on false premises or deceitful persuasion as wrong. The novelty of their argument is to show that rational persuasion can lead to harmful influence; therefore, it should not be taken to be universally a good thing. Another novelty is in showing how the online world is particularly predisposed to turning rational persuasion into wrongful influence due to its specific wrongful features that tend to make rational persuasion into harmful influence: “size (vast quantity of online content); connectivity (access to enormous numbers of users); speed (relatively quick access to that content and those users); precision (efficiency in finding just the content or users that one wants); and disinhibition (users’ willingness to post online what they would not say in person).” (Mitchell & Douglas, 2024, p. 22). Mainstream epistemic concerns with social media platforms tend to focus on manipulation and misinformation, i.e., arational persuasion. Nevertheless, showing that rational persuasion can also be harmful is an essential finding for epistemology; showing that the wrongful features of rational persuasion happen most frequently in online instances of persuasion is important for philosophers of digital technologies.

Mitchell and Douglas’s argument is correct, and I have no qualms about how they arrived at their conclusion. My concern is with what makes rational persuasion harmful in practical day-to-day argumentation settings and where we should go from there. Given their strict definition adopted for rational persuasion (being presented with reasons, i.e. facts that count in favour of the conclusion aimed for the act of persuasion), one wonders how much of this harmful influence is due to the rational part of persuasion as such, namely the factfulness of reasons, and how much is about other factors, belonging to the pragmatics of the argumentation in a social context. While rationality matters a great deal for philosophers, why should it be the ultimate weighing metric for people trying to make up their minds in day-to-day situations? We need to see that LDL or triglycerides are too narrow a metric to be decisive when deciding on a diet overhaul without an expert weighing in and without taking into account the personal circumstances of the inquirer. We need to extend the discussion from instances of individual persuasion to overall evaluations of the online epistemic environment.

In the cases of evaluating online rational persuasion, one should not look only at the standing of the persuader (as suggested by Mitchell and Douglas) but also at the standing of the persuaded party, namely the person doing the queries online. Should I rely on my reasoning capacities alone to decide on a matter I have no expertise in? In other words, do I have the standing to decide on this very complicated and nuanced issue using my rational capacities alone? Perhaps this is an issue of domain-bounded expertise where no rational account can help me decide and I need some extra epistemic skills, domain-specific. Furthermore, rationality, like critical thinking, is a capacity that depends on the environment in which it is deployed. I cannot remain rational in a sea of irrational people; I cannot remain rational if I am in a disabling environment that denies my capacity

as a rational inquirer (see the discussions on epistemic injustice as oppression); I cannot be rational if I do not have the right coordinates for the rationality in a specific domain of inquiry: LDL cholesterol, HDL and triglycerides are all important markers of health, but deciding on one over the other entails other skills than mere rationality or critical thinking, it entails I know some statistics and epidemiology, which are domain-bound skills and expertise. I cannot be rational in an environment that encourages me to overstate my reasoning capacity, giving me the confidence to decide on my own on such sophisticated issues. The fact is that I do not have the epistemic skills to decide on this matter on my own, and others trying to persuade me on the matter may or may not have them, I have no idea. It follows then that I should not decide on this based on merely factual information presented to me. I should remain agnostic. What is the healthiest diet for me? Whatever my GP recommends, based on their medical expertise and knowledge of my medical history. The debate around LDL/ triglyceride as predictors of health? Not my circus, not my monkeys.

The dangers of harmful persuasion online come from encouraging users to over-trust their own rational capacities and take for granted their general epistemic skills in specific domains of expertise. The harm comes from the user's overconfidence in their general reasoning capacities and from the low awareness of the skills gap: what skills are needed to make up one's mind and what skills one has. In this case, evaluating rationality is a good start to show that rationality is never enough on its own. If the online social environment presented by the main social media platforms does not allow for the grounding of my epistemic skills, it is fundamentally toxic. This means that no matter what opinions I gain online or how true these are, I will gain them in wrong ways. General rationality (a true reason that counts in favour of the conclusion) is not fine-grained enough to justify opinion change in the absence of other epistemic skills. Should we use rationality as a metric when evaluating online persuasion? Probably yes, we want to avoid manipulation and misinformation-based persuasion, but rationality on its own should never be the metric for what counts as persuasion online; rather, it should be seen as a necessary but insufficient condition. The interesting discussion stemming from this finding concerns delimiting the other conditions for an instance of rational persuasion that needs to be fulfilled so that it will not count as wrongful.

**Authors' Contributions** All text was written by Lavinia Marin.

**Funding** This work is part of the research programme Ethics of Socially Disruptive Technologies, which is funded through the Gravitation programme of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant number 024.004.031).

**Data Availability** Not applicable.

## Declarations

**Ethics Approval and Consent to Participate** Not applicable.

**Consent for Publication** Not applicable.

**Competing Interests** The authors declares no competing interests.

## Reference

Mitchell, T., & Douglas, T. (2024). Wrongful Rational Persuasion Online. *Philosophy & Technology*, 37(1), 35. <https://doi.org/10.1007/s13347-024-00725-z>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.