The importance of informative interventions in a wicked environment:

Review of 'Breaking the Social Media Prism: How to Make Our Platforms Less Polarizing ' by Chris Bail

David J. Grüning[12*]

Folco Panizza[13]

Philipp Lorenz-Spreen[14]

1: Prosocial Design Network

2: Heidelberg University

3: IMT School for Advanced Studies

4: Max-Planck Institute for Human Development, Center for Adaptive Rationality

*Correspondence:

David J. Grüning

Heidelberg University

Hauptstraße 47

Heidelberg, 69117

Phone: +49 (0163) 8705625

E-mail: david.gruening@stud.uni-heidelberg.de

(Words main body: 1.636)

Introduction

Lewis Thomas (1983), in his medical history treatise, describes the case of an early twentieth century physician, who was eerily accurate in diagnosing typhoid fever in his patients. Only much later did people realize that the physician himself "was a more effective carrier […] than Typhoid Mary" (1983, p. 22), because he was examining his patients using only his bare hands; conferring the disease to everyone he touched. In this story, the complex environmental conditions of disease transmission hid the true causal relationships and led to erroneous conclusions. Thomas had wanted to make a point with his example: such *wicked* environments, in which things are not as they seem, can greatly impair people's ability to correctly infer causes. In the physician's case, people mistook his medical carelessness for diagnostic genius.

Political scientist Chris Bail (2021) provides social media as a more modern example of wicked environments in his book "Breaking the Social Media Prism," where mechanisms are hidden and misperceptions are ample. Many factors contribute to a context's wickedness. These include delayed feedback from decisions (e.g., friend suggestions and requests) and the anonymity of actors (e.g., anonymous or fake user profiles). The most prominent factor for social media is its low traceability of regulating processes, that is, opaque algorithms regulating the presentation of content to users, which is tailored to their interests.

Bail suggests that social media's environmental complexity fuels a severe perception bias. The author regards social media as a *prism,* one that, only crookedly, reflects the true distribution of people's views on different topics. Therefore, users cannot gauge others 'true motives within the space. It is not people's cognition, but their environment's wickedness, which distorts their perceptions.

Revisiting kind and wicked environments

For their part, and to their credit: Thomas and Bail each point out that people must learn about an environment's properties in order to function within it effectively. They are not alone in this respect. For decades, researchers have known that rational decision-making depends upon

ecological factors. (e.g., Gigerenzer et al., 2011; Gigerenzer & Selten, 2003; Simon, 1956). Their work suggests that different phenomena, once constructed as cognitive biases, were, in truth, cognitively efficient tools for navigating within environmental limitations. Rational choice, their works suggest, depends not only on properties of the individuals making the choices — for instance, their capacity and ability — but properties of the environments those individuals occupy — accessibility and diagnosticity of information — as well.

Decision-making environments vary between two extremes: *kind* and *wicked*. Kind environments provide feedback that is swift, accurate, and direct; whereas, wicked environments are based upon opaque variables, which are each intertwined and subject to changes across time. As a result, any feedback in the latter environment is quite volatile, often presented with a time delay (Hogarth, 2001). Kind environments, like board games and most sports, promote a positive learning curve, even for the uninformed decision-maker, because they offer instructive feedback on what should be done next. By repeatedly taking action, people can reliably learn about the environment's conditions and adapt to them. In contrast, wicked environments, like the stock market or social media, can make people spiral into grave misconceptions about their environment and mislead them toward wrong judgements and decisions (Denrell & March, 2001; Einhorn & Hogarth, 1978; Feiler et al., 2012; Koehler & Mercer, 2009).

Behavioral and informative interventions

In his book, Chris Bail presents multiple interventions for the distorted prism. We propose that these can be categorized into two distinct types.

The first kind that he addresses we class as *behavioral interventions*; interventions for behavioral change. Such behavioral interventions are abundantly covered in literature for many decision environments (see e.g., Daley et al., 2018; Glanz et al., 2017; Pennycook et al., 2021; Voelkel et al., 2021). At present, nudges are arguably this intervention type's most prevalent example. Their goal is to *push* individuals towards a specific action. This intervention's applications are wide reaching. In a social media context, Twitter warns about sharing a link without having

accessed it to reduce the distribution of shared misinformation. Recently, Pennycook and colleagues (2021) also proposed that making people pay attention to the accuracy of information reduces the spreading of misinformation.

Bail further presents what we class as *informative interventions*. These interventions provide people the information needed to better understand their environment's features, thereby increasing its transparency. One example of this intervention type is displaying to users the agreement of other readers on a Twitter-post (https://www.prosocialdesign.org/library/highlight-consensus-posts). The feature allows users to get an insight into the consensus of opinions on a respective post. Another instance is Guess et al.'s (2020) intervention for digital media literacy. The authors suggest presenting social media users with tips on where to look for additional online information about news distributed on social media platforms (https://www.prosocialdesign.org/library/list-tips-for-checking-accuracy-of-shared-headlines).

Breaking wickedness with information

Hogarth, Lejarraga, and Soyer (2015) propose that, in general, kind environments are a *necessary* condition for accurate judgments. They conclude that this type of environment should, therefore, be deliberately created for decision-making. We propose an alternative in cases where reshaping the environment from wicked to kind is not possible yet, as it currently is the case for social media. Dissecting the wickedness of this environment by using informative interventions may be sufficient for people to navigate it.

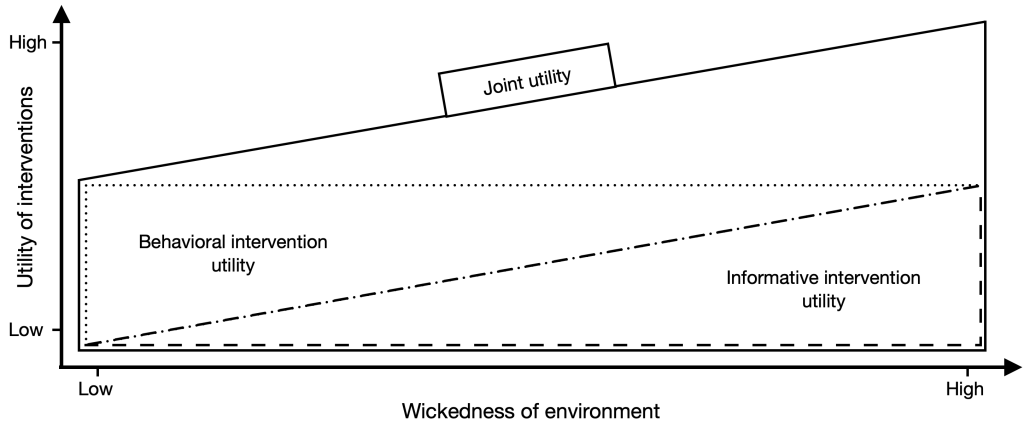
Exclusively applying behavioral interventions in wicked environments may not only prove a wasteful endeavor but also backfire. To explain with an example: A recently discussed behavioral intervention, proposed by Pennycook et al. (2021), is to prompt people to pay attention to news articles' accuracy. However, this intervention may not help or even do harm if the indicators people use to judge an information piece's accuracy are non-diagnostic - for instance, when social indicators are distorted. It is, hence, essential for any given accuracy nudge's effectiveness that an informative intervention accompanies it. These interventions give people the opportunity to

understand the wickedness of the indicators of accuracy and, thereby, make the environment of information and misinformation more navigable.

We conclude that informative interventions are needed within wicked environments in order for behavioral interventions to work as intended. As shown in Figure 1, only the combination of both intervention types can enable people to navigate wicked environments effectively.

Figure 1

Sketch of the utility of behavioral and informative interventions in exclusive and joined formats as a function of the environment's wickedness.



An informative toolbox

Bail presents informative interventions that build on each other as consecutive steps. One of his examples could be used as the beginning of an informative toolbox.

The first of the two steps that Bail emphasizes is granting to people an intuitive, non-judgmental sense of their ideological opinions within a broader political context. The author, on his lab's webpage *Polarization Lab* (www.polarizationlab.com), provides social media users with a web application which they may use to gain insights about how the opinions of journalists, politicians, activists, and other opinion leaders on Twitter relate to one another; and to realize that the opinions of Republicans and Democrats in general are much closer than the wicked social media environment may lead them to believe.

As for the second of the two steps, Bail suggests that we should understand our own position with respect to these opinionated networks. To do this, the *Polarization Lab* offers a self-

report questionnaire, in which users can have their results compared with the general distribution of all other participants. Users may then more accurately comprehend where their views lie within the broader spectrum of opinions on what are often Twitter's most emotionally charged topics. The intervention helps people to see through the distorted view of the opinion landscape that social media leaves them with, where majorities look different than they actually are.

No one intervention in isolation can solve a complex environment's opaqueness. That intervention, however, can be one tool among many. Interweaving multiple informative interventions in a toolbox can jointly address the several caveats for misconceptions in wicked environments, like social media networks, and help behavioral interventions (e.g., the accuracy nudge) to fully unfold.

Conclusion

Chris Bail's diagnosis of the major social media platforms, in his book "Breaking the Social Media Prism," is correct. Social media platforms are wicked environments that promote distorted maps of where users stand in relation to others 'opinions. We, propose that a more detailed distinction between behavioral and informative interventions could benefit Bail's proposed and future solutions.

Furthermore, we invite the wider scientific community to consider a more granular inquiry into—and categorization of—interventions for accurate decision making and its constituent patterns. For example, research on *boosts* as a category of interventions points to a related notion of informative interventions (e.g., Hertwig & Grüne-Yanhoff, 2017; Lorenz-Spreen et al., 2021). Rather than merely highlighting a preferred behavioral pattern, scientists in our field should foster people's inner competencies for effective self-directed action. A successful intervention can be read as one that strengthens the addressee's ability to navigate an environment for themselves.

Advancing research beyond behavioural tools (see e.g., Daley et al., 2018; Glanz et al., 2017; Pennycook et al., 2021; Voelkel et al., 2021) is a promising direction of future interventions for the wicked online environments of our time. Many practical applications of the informative

intervention type for the online environment can already be taken from Bail and new ones can be inspired by them. Their underlying mechanisms can further be blueprints for informative interventions in many other decision-making environments than social media. Collecting information about the context in wicked environments should be a definite step before taking action.

Acknowledgements

We thank John Fallot and Joel Putnam for insightful discussions about the topic and their feedback on an earlier version of the manuscript.

References

Bail, C. (2021). *Breaking the Social Media Prism.* Princeton University Press.

Daley, M. F., Narwaney, K. J., Shoup, J. A., Wagner, N. M., & Glanz, J. M. (2018). Addressing parents' vaccine concerns: a randomized trial of a social media intervention. *American journal of preventive medicine*, *55*(1), 44-54. https://doi.org/10.1016/j.amepre.2018.04.010

Denrell, J., & March, J. G. (2001). Adaptation as information restriction: The hot stove effect. *Organization Science, 12*(5), 523-538. https://doi.org/10.1287/orsc.12.5.523.10092

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*(5), 395–416. https://doi.org/10.1037/0033-295X.85.5.395

Feiler, D. C., Tong, J. D., & Larrick, R. P. (2012). Biased judgment in censored environments. *Management Science, 59*(3), 573-591. https://doi.org/10.1287/mnsc.1120.1612

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199744282.001.0001

Gigerenzer, G., & Selten, R. (Eds.). (2002). *Bounded rationality: The adaptive toolbox.* MIT press.

Glanz, J. M., Wagner, N. M., Narwaney, K. J., Kraus, C. R., Shoup, J. A., Xu, S., O'Leary, S. T., Omer, S. B., Gleason, K. S., & Daley, M. F. (2017). Web-based social media intervention to increase vaccine acceptance: A randomized controlled trial. *Pediatrics, 140*(6), 1–9. https://doi.org/10.1542/peds.2017-1117

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, *117*(27), 15536-15545. https://doi.org/10.1073/pnas.1920498117

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science, 12*(6), 973–986. https://doi.org/10.1177/1745691617702496

Hogarth, R. M. (2001). *Educating intuition*. The University of Chicago Press.

Hogarth, R. M., Lejarraga, T., & Soyer, E. (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science, 24*(5), 379–385. https://doi.org/10.1177/0963721415591878

Koehler, J. J., & Mercer, M. (2009). Selection neglect in mutual fund advertisements. *Management Science, 55*(7), 1107-1121. https://doi.org/10.1287/mnsc.1090.1013

Lorenz-Spreen, P., Geers, M., Pachur, T., Hertwig, R., Lewandowsky, S., & Herzog, S. (2021). Boosting people's ability to detect microtargeted advertising. *Scientific Reports, 11*(15541)**,** 15541. https://doi.org/10.1038/s41598-021-94796-z

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590-595. https://doi.org/10.1038/s41586-021-03344-2

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review, 63*(2), 129-138. https://doi.org/10.1037/h0042769

Thomas, L. (1983). *The youngest science: Notes of a medicine watcher.* Viking.

Voelkel, J. G., Chu, J., Stagnaro, M., Mernyk, J. S., Redekopp, C., Pink, S. L., Druckman, J. N., Rand, D. G., & Willer, R. (2021, May 27). Interventions Reducing Affective Polarization Do Not Improve Anti-Democratic Attitudes. https://doi.org/10.31219/osf.io/7evmp