

Science & Society

Behavioral science should start by assuming people are reasonable



Jens Koed Madsen^{1,8},
Lee de-Wit^{2,8,*}, Peter Ayton³,
Cameron Brick^{4,5},
Laura de-Molier⁶, and
Carla J. Groom⁷

Should policymaking assume humans are irrational? Using empirical, theoretical, and philosophical arguments, we suggest a more useful frame is that human behavior is reasonable. Through identifying goals and systemic factors shaping behavior, we suggest that assuming people are reasonable enables behavioral science to be more effective in shaping public policy.

How policymakers think about human decision-making influences the solutions they propose to challenges like climate change or pandemics. Inaccurate assumptions can result in interventions that are either ineffective or have unintended consequences. While some argue humans are predictably irrational given consistent deviations from the perfectly rational *Homo economicus* [1], there is a long history of criticism such as Lopes' 'rhetoric of irrationality' [2]. Increasingly, behavioral scientists have questioned the utility of using a 'list of biases' to explain behavior [3]. Starting from an assumption of irrationality also risks neglecting contextual factors that could make decision-making adaptive under specific ecological conditions [4].

We argue that behavioral science should start from an assumption that humans are reasonable. We take reasonable to

mean that people with knowledge of the situation (including social pressures) would be able to see the behavior as a satisfactory way to achieve a particular goal. This stance encourages participation from the people subject to these applications and necessitates methodological frameworks that incorporate context. After setting out this conceptual idea, we illustrate how this framework provides a better lens to understand behavioral responses to climate change and coronavirus disease 2019 (COVID-19).

Homo irrationalis?

Claims of irrationality require a normative benchmark against which behavior is evaluated. Perfect information and rational utility optimization are just one possibility. Felin *et al.* [5] argue this benchmark is often contingent on a philosophically unjustified conception of optimality. Across the psychological sciences, they argue that human behavior is often judged against standards that are implicitly assumed to be objective and critique this approach as reflecting an 'all-seeing eye' that can objectively differentiate rational from non-rational choices. Instead, scientists might ask what people are attempting to achieve given their context and environmental constraints. Some decision-making errors can be labeled irrational when one makes explicit the values that underpin what is optimal, like wanting to win more money whilst accepting a bet one knows is bad. However, behaviors that might initially seem irrational may be modeled as optimal use of information available to people within the task [6]. More generally, experimental tasks often make use of artificial scenarios that underestimate people's ability to reason and learn in everyday situations [7]. For example, the adoption of beliefs commonly regarded as irrational can be modeled as a reasonable consequence of how experiences in different social environments shape who can be trusted. Our argument builds on a tradition from scholars such as Simon, Gigerenzer, Lopes, and

others. What is new is our argument that researchers and policymakers need to exhaust alternative explanations based on people's goals and social environments before assuming irrationality and to use methods that engage participants as partners not just targets of interventions.

Homo irrationalis is often accompanied by a seductive vocabulary of cognitive biases [4] offering the appearance of a comprehensive explanation which can be appealing in trying to solve complex policy problems (Box 1). However, this caricature risks being used as a framework to explain behavior in hindsight by retrospectively applying a bias and thereby confecting a misleading impression of explanatory power. This is illustrated by the fact that different biases might point in contradictory directions, such as 'optimism bias' and 'negativity bias' or 'recency bias' versus 'anchoring bias'. It is not clear in advance which of the pairs of biases will manifest in a given situation. Although there is some promising work in unifying bias frameworks, hypothesized biases are often descriptions of an observed behavior, not the demonstration of an internal cognitive process that would generalize to all situations.

Irrational and reasonable responses to COVID-19

The immediate response to COVID-19, particularly in the UK, illustrates how assumptions focused on human irrationality can narrow and compromise behavioral public policy. In the early phases of the pandemic in the UK, decision makers were cautioned that people might not comply with a strict lockdown for longer periods due to 'behavioral fatigue'. Due to concern about the influence of this idea within government, behavioral scientists wrote an open letter questioning the notion of behavioral fatigue, highlighting that while it was plausible that people would tire of strict lockdown rules, one should not assume this single tendency would override

Box 1. Why behavioral science should emphasize reasonableness rather than irrationality

Predicting human behavior is challenging. The explanatory power of behavioral science risks being oversold through the vocabulary of irrationality and bias, especially given the limited estimates of the impact of interventions [14]. Pallesen and Pedersen [15] document how models of human fallibility can dominate the translation from research to policy. Factors like social norms can contribute to public policy, but framing the utility of factors like these in terms of deviations from rationality offers a false allure of objectivity. Assuming reasonableness does not imply that all behavior is ethical or socially acceptable, but it clarifies that deciding what behavioral outcome a policy maker deems to be optimal is a choice. We do not precisely define reasonableness, but it is an advance to acknowledge this challenge rather than relying on a flawed conception of irrationality. The behavioral sciences need to discuss methods to articulate definitions of reasonableness and methods of generating predictions to straddle the gap between *a priori* normative assumptions and complete relativism. This may yield normative benchmarks such as Bayesian updating functions that may be implemented differently from society to society given socio-cultural differences in, for example, who can be trusted as an information source.

people's ability to respond to the exceptional circumstances. Several authors linked the UK government's decision to delay lockdown to this over-extension of the idea of behavioral fatigue [8].

While some broke lockdown rules, there was broad acceptance of lockdown measures despite damage to personal and societal aspects of life, including the economy, social interactions, and mental health. Lockdown compliance and high vaccine uptake both highlight that most people could make substantive changes to their behavior.

Resistance to lockdown rules and vaccine uptake is not explained well by human irrationality. Individuals and countries that adhered more to COVID-19 lockdowns and vaccination acceptance also trusted health authorities more [9]. Vaccine acceptors also perceived more danger, vaccine efficacy, and virulence. In the UK, differences in vaccine acceptance between ethnic groups correlated with prior experience and trust in healthcare systems. These relationships indicate a reasonable logic to lockdown compliance and even to vaccine hesitancy. Assuming people are reasonable refocuses behavioral science on how trust in institutions can be earned and maintained.

Irrational and reasonable responses to climate change

The potentially catastrophic impacts of climate change could suggest that rational

people would rapidly reduce emissions to avoid these consequences. Inaction could therefore be taken as evidence of individual deficiencies or biases. Indeed, this inaction is linked to present-bias. Present-bias describes people's reluctance to sacrifice today for future benefits and is argued to have evolved as a feature of human decision-making to survive in historically unstable environments.

However, assuming people are reasonable moves the focus to different questions and solutions. To illustrate, in the USA people differ dramatically in their assessment of the threats posed by climate change. A 2023 survey by the Pew Research Centre found that 78% of Democrats saw climate change as a major threat compared with 23% of Republicans¹. Such differences point to a lack of political acknowledgement of climate change, not universal cognitive limitations. Atkinson and Jacquet [10] have highlighted how cognitive limitations have been overemphasized as drivers of climate inaction in the popular press and academic literature. To understand the politicization of climate change in the USA, we need to recognize that individuals filter information in alignment with their existing ideas and party allegiances. This does not necessarily mean they are irrational. There are conditions under which confirmation bias can be adaptive [11] and relying on source cues for who one can trust is not fundamentally irrational.

Starting from the assumption that humans are reasonable, the question becomes why Republican politicians became advocates of climate denial and climate inaction and why Republicans are less likely to trust scientists. We believe there are plausible political reasons for this, such as corporate misinformation with regards to climate change [12] and the funding of organizations countering action on climate change [13]. This focus on corporate lobbying as a key factor for behavioral science in the context of climate change is not novel; our analysis reinforces calls made by others that our response to climate change needs to learn lessons from the movement to reduce lobbying from the tobacco industry in reducing harmful smoking behavior.

Attempts to address climate change by targeting individual behavior through taxes on fossil fuels or restricting car traffic have already generated significant political opposition. To dismiss this opposition as irrational risks provoking further polarization around climate action and neglects legitimate concerns people might have. Across Europe, for example, people were previously told by governments to switch to diesel cars because they were better for the environment, but this advice was later reversed. It is therefore reasonable that people would not immediately trust the steps governments are taking. This highlights the importance for governments to develop credible and trustworthy plans that encourage individual behaviors that tackle climate change and increase support for required policies. Assuming people are reasonable makes the communication of credible plans a primary aim of the psychological and behavioral response. This is needed to tackle climate change and to minimize political polarization from more paternalistic actions.

Implications for behavioral science

Assuming that people are reasonable highlights the need to engage with people

Box 2. Participatory methods, from citizens assemblies to co-design, to nudge plus

There are a range of methods that enable more participatory engagement at different stages of an intervention. Prior to launching an intervention, deliberative research ('citizens assemblies') offers the opportunity for structured dialogues with representative cross-sections of the population, enabling people to engage with evidence, experts, and other citizens with whom they may disagree, to reach a deliberative conclusion. Structured discussions like this played a key role in the UK's adoption of a new default enrolment scheme for workplace pensions.

Interventions can also be explicitly co-designed with potential participants to build on social learning and culturally specific knowledge held by participants, such as consulting adolescents in anti-bullying interventions. Interventions can also make use of key community members as the agents of the intervention, for example, to help spread information or shape new norms.

Even in more traditional interventions participants can be encouraged to reflect on whether the suggested change in behavior aligns with their values or goals. This 'nudge plus' approach has the potential to develop a more cyclical process such that citizens are encouraged to feedback to policy makers about the way in which they would like choice environments that shape their behavior to be structured.

who are the targets of interventions. This turns them from targets into partners and also acknowledges the systems they inhabit. This more participatory and deliberative engagement can manifest in methods such as qualitative interviews, focus groups, citizens assemblies, engagement with citizen science initiatives, and deliberative prompts ('nudge plus') during interventions (Box 2).

The medical sciences in the UK now mandate patient and public involvement in all research involving a patient population. Whilst many behavior change projects already involve participatory methods, we think it is time for behavioral science to also adopt formal expectations for participatory engagement in behavior change initiatives, grounded in the assumption that people are reasonable.

Acknowledgments

J.K.M. and L.d-W. were supported by a World Templeton Charity Foundation Grant. C.G. contributed to the paper from a technical perspective. Commentary in the paper on matters of Government policy are not the views of this author or the Department for Work &

Pensions. We would also like to thank Malte Dewies for critical feedback on previous drafts.

Declaration of interests

The authors declare no competing interests.

Resources

www.pewresearch.org/short-read/2023/08/09/what-the-data-says-about-americans-views-of-climate-change/

¹Department of Psychological and Behavioural Sciences, London School of Economics and Political Science, Houghton Street, WC2A 2AE, London, UK

²Department of Psychology, University of Cambridge, Downing Street, CB2, 3EB, Cambridge, UK

³Centre for Decision Research, University of Leeds, Woodhouse Lane, LS2 9JT, Leeds, UK

⁴Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129, 1001, NK, Amsterdam, The Netherlands

⁵Inland Norway University of Applied Sciences, Holsetgata 31, 2318 Hamar, Norway

⁶Decision-context, Am Riedlerberg 5, 83708 Kreuth, Germany

⁷UK Department for Work & Pensions, Human-Centred Design Science Division, Caxton House, Tothill St, London SW1H 9NA, UK

⁸These authors are joint first authors

*Correspondence:

lhd26@cam.ac.uk (L. de-Wit).

<https://doi.org/10.1016/j.tics.2024.04.010>

© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

References

- Ariely, D. and Jones, S. (2008) *Predictably Irrational*, HarperCollins
- Lopes, L.L. (1991) The rhetoric of irrationality. *Theory Psychol.* 1, 65–82
- Hallsworth, M. (2023) A manifesto for applying behavioural science. *Nat. Hum. Behav.* 7, 310–322
- Gigerenzer, G. (2018) The bias in behavioral economics. *Rev. Behav. Econ.* 5, 303–336
- Felin, T. et al. (2017) Rationality, perception, and the all-seeing eye. *Psychon. Bull. Rev.* 24, 1040–1059
- Oaksford, M. and Chater, N. (1994) A rational analysis of the selection task as optimal data selection. *Psychol. Rev.* 101, 608
- Lejarraga, T. and Hertwig, R. (2021) How experimental methods shaped views on human competence and rationality. *Psychol. Bull.* 147, 535
- Feitsma, J. and Whitehead, M. (2022) Behavioural expertise: drift, thrift and shift under COVID-19. *Int. Rev. Public Policy* 4, 149–170
- Lindholt, M.F. et al. (2021) Public acceptance of COVID-19 vaccines: cross-national evidence on levels and individual-level predictors using observational data. *BMJ Open* 11, e048172
- Atkinson, Q.D. and Jacquet, J. (2022) Challenging the idea that humans are not designed to solve climate change. *Perspect. Psychol. Sci.* 17, 619–630
- Rollwage, M. and Fleming, S.M. (2021) Confirmation bias is adaptive when coupled with efficient metacognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 376, 20200131
- Supran, G. and Oreskes, N. (2021) Rhetoric and frame analysis of ExxonMobil's climate change communications. *One Earth* 4, 696–719
- Brulle, R.J. (2014) Institutionalizing delay: foundation funding and the creation of US climate change counter-movement organizations. *Clim. Chang.* 122, 681–694
- Maier, M. et al. (2022) No evidence for nudging after adjusting for publication bias. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2200300119
- Pallesen, T. and Pedersen, K.Z. (2023) Model of human fallibility: traveling behavioral assumptions in public governance. *Perspect. Public Manag. Gov.* 6, 119–130