



# Algorithmic Political Bias—an Entrenchment Concern

Ulrik Franke<sup>1,2</sup>

Received: 1 July 2022 / Accepted: 2 July 2022 / Published online: 1 August 2022  
© The Author(s) 2022

## Abstract

This short commentary on Peters (*Philosophy & Technology* 35, 2022) identifies the entrenchment of political positions as one additional concern related to algorithmic political bias, beyond those identified by Peters. First, it is observed that the political positions detected and predicted by algorithms are typically contingent and largely explained by “political tribalism”, as argued by Brennan (2016). Second, following Hacking (1999), the social construction of political identities is analyzed and it is concluded that algorithmic political bias can contribute to such identities. Third, following Nozick (1989), it is argued that purist political positions may stand in the way of the pursuit of all worthy values and goals to be pursued in the political realm and that to the extent that algorithmic political bias entrenches political positions, it also hinders this healthy “zigzag of politics”.

## 1 Introduction

Peters (2022) explores how algorithmic bias with respect to political orientation can arise and persist, and argues persuasively that this poses substantial and distinctive risks. In particular, Peters identifies scholarships, employments, and academic manuscripts as areas where algorithmic political bias may be at the same time similarly harmful but harder to detect and eradicate than other more well-known algorithmic biases. (For some more general perspectives on algorithmic bias, see, e.g., Mittelstadt et al., 2016, Zerilli et al., 2019, and Franke, 2022—for a synthesis of the literature, see Fazelpour and Danks, 2021.)

The purpose of this short commentary is to present one more distinctive risk related to algorithmic political bias: the risk that such bias exacerbates political entrenchment to the detriment of the polity. This risk is closely related to those originally identified by Peters (2022), and is intended as a complement to these.

---

✉ Ulrik Franke  
ulrik.franke@ri.se

<sup>1</sup> RISE Research Institutes of Sweden, SE-164 29, Kista, Sweden

<sup>2</sup> KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden

## 2 Contingency of Political Positions

At the core of the article by Peters (2022) is the empirical fact that (at least) two political positions, “namely the liberal or politically left-wing viewpoint, and the conservative or politically right-wing orientation” (Peters, 2022, p. 4) are readily observable and detectable in practice. Brennan describes this phenomenon as follows (in a US context):

Consider the following topics: gun control, global warming, how to handle the Islamic State of Iraq and Syria, mandatory paid maternity leave for women, the minimum wage, gay marriage, the Common Core curriculum, and flag burning. If I know your stance on any one of these issues, I can predict with a high degree of reliability what your stance is on all the others. (Brennan, 2016, p. 41)

It is largely because political stances cluster this way that positions can so easily be inferred by algorithms (or fellow humans) and used as predictors in decisions about scholarships, employments, and academic manuscripts. It is also because of this clustering that it is apt to talk about political positions as social identities—they are not just opinions, but opinions which form an identity. This identity aspect is also why Peters (2022, end of Section 4.4) argues that political positions are different from hair color, baldness, or wearing glasses. Without the clustering of (mere) stances into (more deeply felt) social identities, the issue of algorithmic political bias would not have what Peters calls “ethically and epistemically important implications” (Peters, 2022, p. 17).

Now, why does such political position clustering occur? One possibility is that the stances are derived from logically consistent normative political theories. In some circumstances, this may indeed be the best explanation. For example, when Bourget and Chalmers (2014) report from their survey of 931 professional philosophers that 34.8% lean towards or accept egalitarianism, 14.3% lean towards or accept communitarianism, and 9.9% lean towards or accept libertarianism, it is not unreasonable to expect particular stances of the respondents to be logically derived from, or at the very least compatible with, these theories. But professional philosophers are a very particular group. For the general public, this is most probably *not* the best explanation of why clustering happens. A better explanation, argues Brennan, is political tribalism. Even though issues such as those listed in the quote above are logically unrelated, stances on them go hand in hand. “So a good part of the explanation seems to be tribalism: the tribes have settled on answers, and people express fidelity to their tribe by adopting its beliefs.” (Brennan, 2016, p. 42)

This analysis is supported by empirical studies. For example, Cohen (2003) shows in a series of experiments that attitudes towards policies depend almost entirely on the stated position of subjects’ political parties, not on the objective policy content or on what could be inferred from ideological beliefs. Westen et al. (2006) use functional neuroimaging to document how politically partisan subjects use motivated reasoning to avoid negative conclusions about their own

candidates. Brennan offers many additional references throughout chapter 2, pp. 23–54, and Peters (2022, Section 4.4) also notes the prevalence of US political tribalism.

Brennan offers one particularly illuminating observation on the clustering of political stances. Adherents of any one position—US Democrats, say—might claim that the reason why they hold this position is that all the stances on particular issues that it is made up from are simply the right ones, even if they are logically independent.

But even if that were so, why then would Republicans hold the opposite set of beliefs? If Democrats were just unusually good at discovering the truth, that would explain why Democrats converge on one set of logically unrelated beliefs, but it wouldn't explain why Republicans (or non-Democrats in general) converge on the opposite beliefs. We'd instead expect that Republicans would tend to have randomly distributed and disparate beliefs about most of these topics. We'd expect Democrats' beliefs to be positively correlated with one another, but Republicans' beliefs would have few or no positive correlations. We'd expect Democrats' beliefs to form a cluster, but not Republicans' beliefs. (Brennan, 2016, p. 42)

Thus, while the political positions readily detected by algorithms are empirically stable, these social identities are at the same time logically contingent. They could have looked very differently. We now proceed to look further at the construction of these identities, and the role played by algorithms in this process.

### 3 Construction of Identities

What does it mean that political positions are constructed social identities? It is obvious that political positions are social in the sense that they concern society—its political and economic organization, the relations between its citizens, etc. But the construction of political positions as social identities refers to something else; to “the *matrix* within which an idea, a concept or kind, is formed” (Hacking, 1999, p. 10, emphasis in original) and to how such matrices can affect how people are perceived by themselves and others.

Reusing an example of women refugees in Canada from Moussa (1992), Hacking (1999) explains how this works:

When we read of the social construction of *X*, it is very commonly the idea of *X* (in its matrix) that is meant. And ideas, thus understood, do matter. It can really matter to someone to be classified as a women refugee; if she is not thus classified, she may be deported, or go into hiding, or marry to gain citizenship. The matrix can affect an individual woman. She needs to become a women refugee in order to stay in Canada; she learns what characteristics to establish, knows how to live her life. By living that life, she evolves, becomes a certain kind of person (a women refugee). And so it may make sense to say that the

very individuals and their experiences are constructed within the matrix surrounding the classification “women refugees.” (Hacking, 1999, p. 11)

Now, in our context, much the same could be said about political positions *qua* social identities. Psychological and social mechanisms discussed in the previous section are part of the matrices of political social identities; part of the matrices surrounding classifications such as republican, democrat, conservative, liberal, socialist, and libertarian. And furthermore, Peters’s concern can be described as the concern that algorithms are becoming part of these matrices as well, since it may no longer be possible to hide political views (Peters, 2022, see especially Section 4.3). In order to get scholarships or employments, or to get academic manuscripts published, we may, to paraphrase Hacking ‘learn what characteristics to establish, know how to live our lives’ to fit the ideas of certain political positions. Of course, we do not claim that algorithms make up the entire matrices. Political tribes—and more benign political identities—predate modern AI technology. But, following Peters’s arguments, algorithms are certainly becoming important parts of the matrices surrounding political social identities.

#### 4 Zigzag, Change and Development

We have argued first that the sets of stances that make up political positions are typically contingent, and second that algorithmic processing nowadays make up part of the matrices which affect—and reify—the corresponding identities. But what precisely is the problem with these political social identities?

In the essay *The Zigzag of Politics*, Nozick (1989, pp. 286–296) argues that swing voters give electoral democracies a mechanism to pursue multiple values in the political realm. All values and goals which deserve to be pursued cannot “be pursued with full energy and means, and perhaps these goals are theoretically unreconcilable also, in that not all good things can be adjusted together into a harmonious package” (Nozick, 1989, p. 292). But even if they cannot be pursued at the same time, they can be pursued one after another:

The electorate wants the zigzag. Sensible folk, they realize that *no* political position will adequately include all of the values and goals one wants pursued in the political realm, so these will have to take turns. The electorate as a whole behaves in this sensible fashion, even if significant numbers of people stay committed to their previous goals and favorite programs come what may. For there may be a significant swing bloc of voters that will shift to new goals and make the difference (Nozick, 1989, p. 295, emphasis in original)

The electorate perceives the need for change quicker than their parties and elected politicians, whose positions are more entrenched. These latter will see no need to shift positions until “emphatically told to do that by the electorate” (Nozick, 1989, p. 294). In this sense the electorate can be smarter than the politicians.

However, following Nozick’s argument, we can now see that if the entry of algorithms into the matrices of political social identities further entrenches those

identities, this may make the voters more like the politicians. If too many stay committed to their positions come what may, the healthy zigzag will be impeded. Again, as noted in the previous section, it should be stressed that algorithms are not the entire matrices, but they are certainly part of them.

## 5 Conclusions

We have identified the entrenchment of political positions as one additional concern related to algorithmic political bias, beyond those identified by Peters (2022). In Section 2, we observed that the political positions detected and predicted by algorithms are seldom theoretically necessitated. On the contrary, the clusters of stances that make up a political position are typically contingent in their compositions due to the psychological phenomenon called “political tribalism”, as argued by Brennan (2016). In Section 3, we followed Hacking (1999) to analyze the social construction of political identities, concluding that algorithmic political bias can contribute to such identities—and to the tribalism—even though there are certainly other such contributing factors as well. In Section 4, we followed Nozick (1989) to argue that purist political positions may stand in the way of the pursuit of all worthy values and goals to be pursued in the political realm. Indeed, the electorate may be smarter—in the sense of perceiving the need for and voting for the necessary zigzag—than their parties and elected politicians, whose positions are more entrenched. However, algorithmic political bias may make the voters more like the politicians, by constantly reifying and entrenching their political positions.

**Author Contribution** Not applicable (single author)

**Funding** Open access funding provided by RISE Research Institutes of Sweden.

**Availability of Data and Material** Not applicable

## Declarations

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

**Consent for Publication** Yes

**Conflict of Interest** The author declares no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bourget, D., & Chalmers, D.J. (2014). What do philosophers believe? *Philosophical Studies*, 170(3), 465–500. <https://doi.org/10.1007/s11098-013-0259-7>.
- Brennan, J. (2016). *Against democracy*. Princeton: Princeton University Press. <https://doi.org/10.1515/9781400882939>
- Cohen, G.L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, 85 (5), 808. <https://doi.org/10.1037/0022-3514.85.5.808>.
- Fazelpour, S., & Danks, D. (2021). Algorithmic bias: senses, sources, solutions. *Philosophy Compass*, 16(8), e12760. <https://doi.org/10.1111/phc3.12760>.
- Franke, U. (2022). First- and second-level bias in automated decision-making. *Philosophy & Technology*. <https://doi.org/10.1007/s13347-022-00500-y>.
- Hacking, I. (1999). *The social construction of what?* Cambridge: Harvard University Press.
- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679. <https://doi.org/10.1177/2053951716679679>.
- Moussa, H. (1992). *The social construction of women refugees: A journey of discontinuities and continuities*. PhD thesis, University of Toronto.
- Nozick, R. (1989). *The Examined Life: Philosophical Meditations*. Simon and Schuster.
- Peters, U. (2022). Algorithmic political bias in artificial intelligence systems. *Philosophy & Technology* 35. <https://doi.org/10.1007/s13347-022-00512-8>
- Westen, D., Blagov, P.S., Harenski, K., Kiltz, C., & Hamann, S. (2006). Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 US presidential election. *Journal of Cognitive Neuroscience*, 18(11), 1947–1958. <https://doi.org/10.1162/jocn.2006.18.11.1947>.
- Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (2019). Transparency in algorithmic and human decision-making: is there a double standard? *Philosophy & Technology*, 32(4), 661–683. <https://doi.org/10.1007/s13347-018-0330-6>.

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