Who gets to shape the algorithms we use?

Philipp Chapkovski · 12. Oktober 2026
Über Politik aus der Wissenschaft

Recommendation algorithms help determine which news we read and whose opinions we encounter. Research often asks how these systems influence political divisions. We turn the question around: could political divisions influence the information that algorithms learn from?

Politics can reach into seemingly unrelated interests. Analysing millions of online book co-purchases, Shi et al. (2017) found that liberal political books were more closely associated with basic science books, while conservative books were more closely associated with applied science. These were purchasing patterns, not evidence that algorithms caused polarization.

Recommendation systems learn from user behaviour and feedback. But not everyone contributes equally, and the feedback algorithms receive may not represent the preferences of all users (Schnabel et al., 2016). If political divisions influence who provides feedback, they may also influence whose preferences the algorithms learn to represent. But what happens when providing that feedback is costly and also benefits someone you disagree with?

Who pays to improve algorithms?

A shared system may need to balance accuracy with diversity, fairness, or political balance. Users can disagree about those priorities. If feedback helps shape the system, willingness to contribute could influence whose values it reflects.

Our theoretical model suggests that similarity can encourage free-riding. Imagine using the same recommendation system as someone who wants exactly what you want. If they spend time improving it, you benefit too. Why not let them do the work?

But similarity could also encourage cooperation. Previous experiments in the US have found that people cooperate more with politically like-minded partners, even in tasks that have nothing to do with politics (Robbett & Matthews, 2023; Whitt et al., 2021). People may simply be more willing to help those who share their views, and less willing to benefit political opponents. This creates two competing forces: similarity may encourage free-riding, but it may also strengthen cooperation.

A small price for better recommendations

We explored this tension in a small online pilot with 199 UK and US participants recruited through Prolific. In an interactive oTree task, they first provided preference information and ranked five real movies. They then made ten decisions about rating fictional movie profiles, contributing 4.4 ratings on average.

Each started with €2.50. Before each decision, participants saw whether an anonymous partner was described as similar or opposed on political views or everyday preferences. Political matching described the partner, rather than necessarily the movie's content.

Rating cost €0.25; skipping kept that money. The incentives were designed so that a rating added approximately €0.20 to each user's expected bonus: €0.40 together. The private benefit was smaller than the cost, but the joint benefit was larger. That is the public-goods dilemma.

Importantly, every rating increased the expected earnings of both users by the same amount, regardless of whether they agreed politically. This allowed us to isolate the social side of the dilemma: does knowing who else benefits change our willingness to contribute.

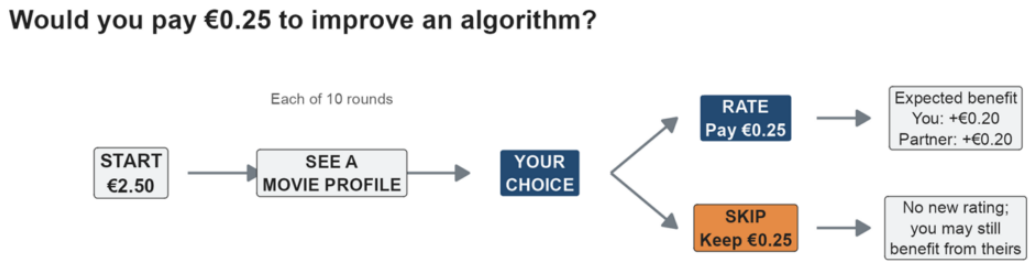


Figure 1. Designed expected returns from one rating, before subtracting its cost. These are approximate expectations, not guaranteed payments.

Two countries, opposite tendencies

Interestingly, the two countries showed opposite patterns. In the UK, participants tended to contribute more when paired with someone holding different preferences, consistent with our free-riding prediction. In the US, however, participants tended to contribute more when their partner was similar to them. Although these differences are statistically uncertain, the contrasting patterns offer an interesting starting point for further investigation.

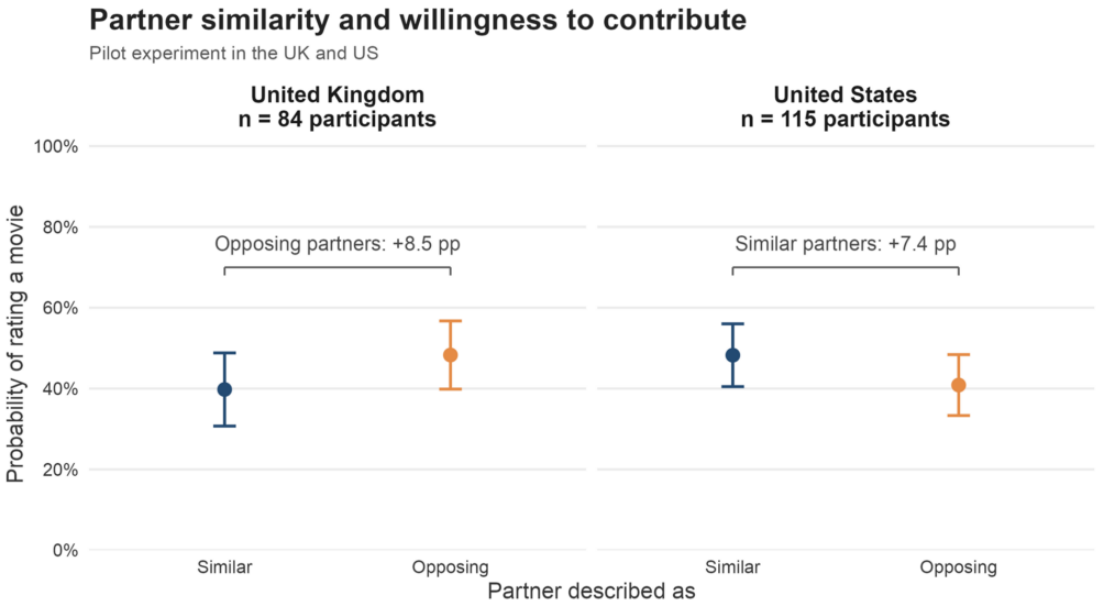


Figure 2. Willingness to provide costly feedback when paired with similar or opposing partners in the UK and US. Labels show the observed differences in percentage points; error bars represent 95% confidence intervals calculated at the participant level. The

patterns are exploratory and the differences remain statistically uncertain.

Could affective polarization, dislike or distrust across political divides, help explain the pattern? Boxell, Gentzkow and Shapiro (2024) document a sharper long-run increase in the US than in Britain. Whitt et al. (2021) find that US partisans contribute less to a shared public good when paired with political opponents.

Perhaps, in a more polarized society, who benefits from a contribution matters more than the temptation to save money. People may be willing to sacrifice something to help those they identify with, while hesitating when their contribution would also benefit someone they disagree with. Our pilot offers an intriguing hint in this direction, although the evidence is still preliminary. The larger experiment will allow us to investigate whether political divisions really shape these decisions.

What comes next?

We are preparing a larger experiment with 2,400 participants, equally divided between the UK and US. It will test whether these tendencies hold and examine political similarity more closely. Beyond movies, could political divisions shape whose preferences are represented in the algorithms we all use?

Stay up to date: Join our mailing list and get a short email whenever a new post goes live, usually every three weeks. No ads, and you can unsubscribe at any time. Sign up here

References

1. Shi, F., Shi, Y., Dokshin, F. A., Evans, J. A. and Macy, M. W. (2017). Millions of online book co-purchases reveal partisan differences in the consumption of science. *Nature Human Behaviour*, 1, 0079.
2. Schnabel, T., Swaminathan, A., Singh, A., Chandak, N. and Joachims, T. (2016). Recommendations as Treatments: Debiasing Learning and Evaluation. *Proceedings*

of the 33rd International Conference on Machine Learning, PMLR 48, 1670–1679.

3. Boxell, L., Gentzkow, M. and Shapiro, J. M. (2024). Cross-Country Trends in Affective Polarization. *Review of Economics and Statistics*, 106(2), 557–565.
4. Whitt, S. et al. (2021). Tribalism in America: Behavioral Experiments on Affective Polarization in the Trump Era. *Journal of Experimental Political Science*, 8(3), 247–259.
5. Robbett, A. & Matthews, P. H. (2023). Polarization and Group Cooperation. *Quarterly Journal of Political Science* , 18, 215–241.

Philipp Chapkovski (2026): Who gets to shape the algorithms we use?. Empirische Politik UDE.
https://www.politik-wissenschaft.org/2026/10/12/who-gets-to-shape-the-algorithms/
Lesefassung aus dem Onlinebeitrag. Interaktive Inhalte sind in der Onlinefassung verfügbar.