Is the Internet safe? Personalization and the threat of epistemological skepticism
Charles Côté-Bouchard (Carleton University) charles.cotebouchard@carleton.ca

Overall thesis: Although the Web gives us access to more truths than ever, its increasingly *personalized* character also prevents us from *knowing* many of those truths. And since we increasingly rely on personalized platforms for belief formation, this knowledge undermining effect risks affecting a significant portion of our beliefs.

Argument in a nutshell: because of its usage-maximizing aim, personalization introduces an element of epistemic *luck* or *instability* that tends to prevent personalized true beliefs (true beliefs formed via personalized platforms) from constituting knowledge.

Part 1. The Personalized Web

1.1 What personalization does and what it seeks

- What it does: uses algorithms to automatically select what to show to each individual user.
- What aim guides that selection:
    - Final aim: maximizing usage/engagement
    - Instrumental aim: prioritize content that promotes engagement + filter out content that won't
- Personalization: *The use of algorithms, by online platforms, to automatically curate and filter content in a way that is personalized for each individual user, based on user data, and with the aim of maximizing usage.*

1.2 Features of usage-maximizing content

- Attention-grabbing features: Sensational, surprising, interesting, funny, outrageous, simple, fun
    - As opposed to boring - complex, nuanced, tedious)
- Sympathetic features: In adequation with or confirming your worldview and identity
    - As opposed to hostile - embarrassing counterevidence; content calling your identity and worldview into question)

1.3 What it favors (Intermediary conclusion): *Personalization, when it works well, tends to prioritize attention-grabbing and sympathetic content over boring and hostile content*

1.4 It is increasingly part of the etiology of our beliefs

1.5 The Web contains a significant quantity of *misinformation*

- Falsehoods, propaganda, half-truths, misleading truths, 'fake news'
- Increasingly hard to detect:
    - Can look reliable or real (deepfakes)
    - Speaks to our cognitive biases
    - Often second-hand

1.6 A lot of online misinformation is attention-grabbing and sympathetic, often more so than reliable content

1.7 Personalization is susceptible to favor misinformation over reliable content. (Intermediary conclusion)

Part II. Personalization and Epistemology

2.1 Knowledge requires a stable connection with the truth
- Prominent conditions: (i) sensitivity, (ii) safety, (iii) relevant alternatives, and (iv) no-defeat

2.2 Assumption: if a true belief does not meet knowledge conditions (i)-(iv), then it lacks the stable truth-connection required for knowledge.

2.3 Personalization tends to prevent our true online-based beliefs from meeting knowledge conditions (i)-(iv).

- **Situation W:** you form the belief that p, which fits your worldview, via a personalized platform.
i. **Personalization vs. sensitivity**
    a. <u>Sensitivity condition</u>: If p were false, you would not believe that p; you do not believe p in any of the closest possible worlds where p is false.
    b. <u>Sensitivity and personalization in W</u>: Would you continue to believe that p if p were false? Personalization makes that likely. Would still prioritize content that supports your belief that p (including misinformation) over embarrassing counterevidence.
ii. **Personalization vs. safety**
    a. <u>Safety condition</u>: If p were false, you would not believe that p; you do not believe p in any of the closest possible worlds where p is false.
    b. <u>Safety and personalization in W</u>: personalization tends to increase the number of nearby possible worlds in which you believe that p, including, possible worlds in which p is false (since personalization is not truth-tracking).
iii. **Personalization vs. relevant alternatives (RA)**
    a. <u>RA condition</u>: you know that p only if you can rule out all relevant alternatives to p
    b. <u>RA and personalization in W</u>: makes following alternative increasingly relevant and not ruled out:
        i. the personalized platform has given you misinformation about p because it was attention-grabbing and/or sympathetic, and/or it has hidden strong counterevidence because it was less attention-grabbing and/or sympathetic.
iv. **Personalization vs. no-defeater**
    a. <u>No-defeater condition</u>: there is no true proposition such that if it were added to your evidence for p, you would no longer be justified in believing p.
    b. <u>No-defeater and personalization in W</u>: personalization susceptible to introduce the following sort of defeater:
        i. (D) the platform's algorithms may very well have (i) shielded you from reliable counter-evidence because it was less attention-grabbing and in tension with your worldview, and/or (ii) prioritized hard-to-detect misinformation because it was more attention-grabbing and in agreeable with your identity.

Part III. Skeptical conclusions

3.1 Personalization threatens to prevent a significant portion of online-based belief from constituting knowledge. (from 2.1-2.3)

3.2 Since we increasingly rely and depend on personalized platforms, and since our beliefs are highly interconnected, our beliefs *in general* are increasingly vulnerable to the knowledge-undermining effect of personalization.