Towards a Critical Social Epistemology of Social Media

Joshua Habgood-Coote,
DRAFT – please ask before citing

Abstract
What are the proper epistemic aims of social media sites? A great deal of social media critique presupposes an exceptionalist attitude, according to which social media is either uniquely good, or uniquely bad for our collective knowledge-generating practices. Exceptionalism about social media is troublesome, both because it leads to oversimplistic narratives, and because it prevents us making relevant comparisons to other systems. The goal of this chapter is to offer an anti-exceptionalist account of the epistemic aims of social media. I will argue that social systems (and epistemic institutions in general) ought to pursue three distinct epistemic goals: promoting good epistemic outcomes for users, realising epistemically good institutional features, and achieving structural epistemic justice. Although these goals are often mutually supportive, I will consider a number of cases in which these values lead to dilemmas about how to design epistemic institutions, which can only be resolved by appealing to ethical considerations. This discussion highlights that just as in the political realm, there is no such thing as a neutral epistemic system.

Introduction
Discussions of social media—and of the internet in general—tend to work on the assumption that is a world apart which cannot be governed, predicted, or evaluated by the same rules as the rest of our social lives. The optimistic version of exceptionalism emerged through the 90s, leading to Barbrook and Cameron (1995) call the Californian Ideology. This discourse presented social media as the herald of a democratisation of discourse, a renewal of liberal freedoms, and an explosion of dynamic cultural production. Since 2016, the optimistic outlook has soured, but without touching the underlying exceptionalist attitude. Social media has become a unique source of pernicious misinformation, the driver of the algorithmic fracturing of the public sphere, and a cesspit of unifying ‘low’ culture. In the popular imaginary, social media has transformed from a cornucopia of liberal freedoms into a haunted house of anti-democratic threats. While this techno-pessimist narrative undoubtedly gets some important things right, its underlying exceptionalism means that it obscures at least as much as it reveals. This exceptionalist attitude allows social media’s problems to be packaged up in ahistorical neologisms (‘fake news’, ‘echo chambers’, the ‘infodemic’), and makes it easy to present social media as a
distinctive realm with its own ideals, which are often picked to defend the status quo and protect corporate interests.¹

Within philosophical discussions, exceptionalism about social media has manifested itself in a defence of established epistemic systems. The story goes like this: established epistemic systems have positive feature X, social media either undermines X, or replaces it with the inferior feature Y. So, current social media ought to be reformed to closer resemble established epistemic systems.² Single-feature comparisons have a tempting simplicity, but this narrative ignores the problems of established epistemic systems, while obscuring the virtues of social media.

This paper is an experiment in developing an anti-exceptionalist social epistemology of social media. The goal is to treat social media as one among many epistemic systems, to be evaluated according to how it contributes to our collective pursuit of knowledge, understanding, and truth. In line with the methodology of systems-oriented social epistemology (Goldman 2010), we will be evaluating social media as a social system for generating good epistemic outcomes. However, unlike many deployments of systems oriented social epistemology, we will be considering three different kinds of epistemic outcomes of epistemic systems: their consequences for the epistemic states of individuals, their institutional properties, and their contributions to epistemic justice. This pluralist approach highlights a fact that is obscured by single-feature comparisons: the epistemic desiderata for social media platforms (and in fact, of epistemic systems in general) are often in tension with one another and must be negotiated by considering values outwith epistemology.

I will call this pluralist approach to the epistemic evaluation social systems Critical Social Epistemology. This approach considers how political oppression interferes with the promotion of epistemic goods, the realisation of institutional ideals, and the political and moral issues involved in trade-off cases. This nomenclature also signals the connections to between social epistemology and critical projects in technology

¹ In a memo from 2016, Facebook executive Andrew Bosworth defended the platform’s role in data collection and connection to terrorism by presenting ‘connectedness’ as the platform’s sole ideal:

The ugly truth is that we believe in connecting people so deeply that anything that allows us to connect more people more often is *de facto* good. It is perhaps the only area where the metrics do tell the true story as far as we are concerned […] It is literally just what we do. We connect people. Period. (quoted in Mac, Warzel, Kantrowitz 2018)

This memo is part of a longer rhetorical strategy of leveraging the value of social connectivity to motivate the collection of huge amounts of personal data about its users (see van Dijck 2013).

² For an example of this narrative see (Goldman 2008) on the social epistemology of blogs and see (Coady 2011) on the problems with Goldman’s conservatism.
studies, including critical information studies (Noble 2018), critical code studies (Marino 2020), and critical disinformation studies (Marwick, Kuo, Cameron, Weigel 2021).

The plan of action is as follows. The first section fixes the target of analysis by asking what a social media site is. The second section lays out critical social epistemology, unpacking the different epistemic goals which the approach appeals to—individual epistemic states, institutional epistemic properties, and epistemic justice—and elaborates the cases in which these goals come into conflict. The third section applies critical social epistemology to social media, showing how several debates implicitly appeal to the different flavours of epistemic evaluation. This section considers how anti-misinformation nudges manifest a concern with individual epistemic outcomes, how worries about algorithmic polarisation manifest a view about the institutional epistemic ideal for social media, and how concerns about algorithmic bias and technological redlining connect to epistemic oppression. The fourth section shows that the conflicts between these three epistemic goals also show up in debates about social media. We will focus on three conflicts: i) how highly connected networks can benefit individuals while creating institutional problems, ii) how in non-ideal circumstances the pursuit of genuinely good institutional properties (such as a unified public sphere) can worsen epistemic oppression, and iii) the how goal of maximising epistemic goods for individuals can lead to unequal distributions of goods, and the flouting of rights. The fifth section ties things up and makes some speculative comments about what a critical social epistemology of social media might look like.

1. Social Media

What is a social media site? I take it that a social media site is one whose primary service is to allow users to hang out, where the majority of content shared on that site is generated by users. Many sites which are not social media sites host content which is primarily created by users (think of wikipedia, Tinder, and Quora). These sites will not count as social media because although some people may use them to hang out, this is not their primary service. Other sites primarily function to people to hang out, without being social media sites (think of online multiplayer video games). These cases are excluded because their content is not user generated.

Let’s set out some non-site-specific terminology. I will talk about sites having users who will typically have a profile displaying personal information, and will be

---

3 Note: not necessarily users of that platform. Facebook would still be a social media site even if all of its posts came from other sites. For alternative definitions, see (boyd and Ellison 2007, Sunstein 2018, 22, Zuckerman and Rajendra-Nicolucci 2020).


5 This is not to say that every case is clear: online games like Minecraft which allow user-generated content are fringe cases, and it isn’t obvious if we should think about online videoconferencing as a kind of social media.
connected to various other users either through symmetric or asymmetric relationships. Users are able to post content, in the form of pictures, text, video, or audio, and are also able to repost other people’s content. Posts may be accessible to different groups of people, depending on the set-up of the site: it might be that only other connected users are able to view posts, but most sites will combine aspects of public and private communication (for example a public profile, and private messaging). Sites will typically be organized around a timeline on the homepage that organizes posts from other users, typically using some element of algorithmic sorting to determine the order and salience of posts. Insofar as there is something distinctive about social media sites, it is their combination of networking large number of participants with the intimacy, informality, and reliance on personal credibility that would be more normally found in private conversations.

Within this basic framework, social media sites are extraordinarily heterogenous. Different sites have users from different demographics, different communities find different uses for social media—from emergency mycological advice, to LARPing the 1990s—and sites develop over time in ways that are often difficult to track. Users have radically diverse motivation-sets for using social media, many of which have nothing to do with the desire for true beliefs.

From the point of view of systems-oriented social epistemology, a social media site is not merely a piece of technology; it is a socio-technical assemblage that meshes a technological system that has certain communicative affordances with a set of social practices. This mesh goes both ways. The social practices of users will be constrained—but not completely determined by—communications technology, and vice versa for the intended workings of the technological system. This means that the epistemic outcomes depend on both social and technical features of social media: a social media site which allows unlimited reposting might be epistemically disastrous if it is combined with reckless social practices for reposting, whereas the same technology might be of great value if combined with judicious social practices for reposting.

Many critical discussions of social media have focused on one part of this system. Calls for media literacy education focus exclusively on social practices, whereas calls to remove the ‘like’ button (or its analogues) focus on technological affordances. Both kinds of interventions might be useful, but in order to think systematically about the epistemic features of social media, we need to think about both its social and technological features, including norms for posting (Rini 2017), moderation practices and community guidelines (Roberts 2019, Frost-Arnold forthcoming), the affordances of sites (Nguyen forthcoming), the underlying economic model (Castro and Pham 2020), and the epistemic structure of online communities (Nguyen 2020).

2. Critical Social Epistemology

There are several different ways in which one might carve up different approaches in social epistemology. For our purposes, we can distinguish approaches
axiologically. This gives us three approaches to evaluating social systems from: *individualist* approaches which evaluate social systems by considering their effects on the epistemic states of individuals, *institutionalist* approaches which evaluate social systems by considering their institutional epistemic properties, and approaches which focus on *epistemic justice*, considering the ways in which the workings of epistemic systems go wrong in non-ideal cases.

2.1. Individualism

Individualist approaches evaluate social systems by considering their consequences for the people who participate in or rely on the system. Broadly speaking, a system is good if it causes its participants to have good (or improved) epistemic states, and bad if it causes them to have bad (or worsened) epistemic states. Individualist approaches are typically consequentialist, in the sense that social systems are evaluated by considering their epistemic consequences for participants. The best-known individualist approach to social epistemology is Alvin Goldman’s veritism (Goldman 2000). Veritism takes the accuracy of individual belief to be the basic epistemic aim and evaluates social systems by considering whether they make participants’ beliefs on average more or less accurate. There are two parts of the accuracy of belief which we might want to distinguish: the possession of true beliefs, and the avoidance of false beliefs (Coady 2011). Although Goldman weighs these two goals equally—at least in his official theory—we might want to say that having true beliefs (or closeness to the truth) is more valuable than having false beliefs (or closeness to the false) is disvaluable (Pettigrew 2016). Despite its popularity, veritism is only one of a family of individualist approaches to social epistemology. We might look at different epistemic outcomes (such as justification, knowledge, or epistemic virtue), take a different measure of accuracy (replacing average accuracy with total accuracy), add in considerations about the distribution of epistemic goods between people or questions, or offer a non-consequentialist axiology. With this said, veritism will be our exemplary individualist approach.

2.2. Institutionalism

Institutionalist approaches to social epistemology evaluate social systems by considering their institutional properties. For example, one might consider how reliable a consensus position is (Zollman 2007, 2010), how effective a system of scientific inquiry is at reaching truth (Longino 1990, 2002, Kitcher 1990, Strevens 2002S), or the problems caused by polarization (Nguyen 2020). Typically these properties will be at least somewhat independent of the epistemic properties of the users of a system. The reliability of a consensus depends on individual reliability and the aggregation method (List 2005), a well-functioning system of inquiry can be composed of inquirers who care principally about individual enhancement (Zollman

---

6 To my knowledge, there are no worked out non-consequentialist individualist approaches to social epistemology, although there is nothing to stand in the way of a social deontological epistemology or social virtue epistemology.
2018), and polarisation can emerge from a community of rational agents (Pallavichini, Hallsson, and Kappel 2018). Following (Mayo-Wilson, Zollman, and Danks 2011), I will call the idea that the methodological prescriptions for individuals and communities (and hence their epistemic properties) diverge, the independence thesis. There are two threads in institutionalist social epistemology: the use of formal models to think about epistemic communities (see Zollman 2013, Boyer-Kassem, Mayo-Wilson, and Weisberg 2017), and work in feminist epistemology of science, which stresses that phenomena such as disagreement, testimony, and the role of values in inquiry should be assessed by considering their role in collective practices of inquiry (Longino forthcoming). Institutionalism doesn’t have a stand-out example, and we will be mainly relying on results about the communication structure of epistemic systems (Zollman 2007, 2010, forthcoming).

2.3. Epistemic Justice

Individualist and Institutional approaches have a tendency to idealise away the messy details of actual epistemic systems, including the way in which political power and domination interferes with the proper operation of epistemic systems. Epistemic Justice-focused approaches correct this tendency, taking what we might think of as a non-ideal approach to social epistemology. Whereas individualist and institutionalist approaches start by developing an epistemic ideal, before asking what a real-life system which instantiated that ideal might look like, the epistemic justice approach starts by characterising the problems of real-life systems, before asking what ameliorative measures might look like.

This approach to epistemology is extremely diverse, and we will focus on epistemic oppression, the distribution of epistemic goods, and inverted epistemologies.

Epistemic oppression (Fricker 1998, Dotson 2014) refers to the different ways in which individuals can have their ability to contribute to collective inquiry hampered, for example by means of social stereotypes which contribute to them being given less credibility that they deserve (what Fricker calls ‘testimonial injustice’), and by means of conceptual systems which hamper their ability to articulate their social experiences (what Fricker calls ‘hermeneutic injustice’) (Fricker 2007). Concerns about the distribution of epistemic goods arise from the observation that epistemic outcomes can be differentially distributed between different people, groups, and research agendas (Fallis 2003, Coady 2010, 2017). For example, a system for health funding which failed to invest any resources in research for issues that affect people with African heritage would be bad (in part) because it produced an unequal distribution of knowledge across important questions. Inverted epistemologies are epistemic systems which cultivate ignorance which serves the interests of oppressive political systems. The best-known inverted epistemology is white ignorance: the

---

7 Although feminist social epistemology’s engagement with the sociology of science does a lot to correct this tendency.
8 On non-ideal theory in general see (Mills 2005), and on non-ideal theory in epistemology, see (Mckenna forthcoming).
often-obscured social system which functions to produce ignorance which serves a white supremacist social order (Mills 1997, 2007), but there are many other examples (Proctor and Schiebinger 2008). Following Anderson (2012), I will think about epistemic justice as a feature of social systems, which obtains when those systems respect the epistemic agency of relevant individuals, distribute epistemic goods in an equitable way, and do not contribute to inverted epistemologies.  

2.4. Conflicts

Having distinguished the different epistemic aims of social systems, we need to complicate things by showing how these aims can come into conflict with one another. These conflict pair-wise: there are conflicts between individual epistemic goods and institutionally valuable features, conflicts between institutionally valuable features and epistemic justice, and conflicts between maximising individual epistemic goods and epistemic justice.

2.4.1 Individualism and Institutionalism

The independence thesis opens up the possibility of social dilemmas for epistemic goods: cases in which it is epistemically best for each individual to take one option, but every individual taking that option creates a worse outcome for the group. There are several examples of social dilemmas for epistemic goods, (see Kummerfeld and Zollman 2015, Dunn 2018), but we will focus on choices of the communication structure of a system in (here I am relying heavily on Zollman forthcoming).

Consider a model where a group is deciding on a binary factual question by majority vote. Each individual starts with the same prior belief, and then receives different evidence which is equally likely to be correct. They update on this evidence via Bayesian reasoning to reach a view about the answer to the question, before expressing this credence by voting. Each individual desires that their own vote is correct. If none of the individuals in the model are connected, we meet the assumptions of the Condorcet Jury theorem (List and Goodin 2001) meaning that the majority position will be more reliable than the average accuracy of individuals in the group. If we take the same model and introduce social learning, things look rather different. In a roll-call situation in which individuals vote in sequence, and each individual can see all prior votes, the group faces the threat of an information cascade. If the first two voters have received evidence that favours option A, even if the third voter has evidence that favours B, she will vote for A because the evidence

---

9 Two loose ends which we won’t have space to resolve: can the three parts of structural epistemic justice conflict with one another? How should we think about the relation between individual and system-level features of epistemic oppression, and the measures taken to ameliorate each?

10 See also (Bishop 2005), (Smart 2018), (Levy and Alfano 2020).

11 The possibility that information cascades involve free-rider dynamics was first noted by (Banarjee 1992, 816), and is explored by (List and Pettit 2004), and (Dunn 2018) (who argues that these situations ought not to be described as social dilemmas).
from their peers outweighs their private evidence (and so on for all later voters). In this situation the group’s position can be determined by the luck who votes first, meaning that roll-call groups will produce less reliable results perform than unconnected groups.

The choice between the unconnected and roll-call scenarios leads to a social dilemma: individuals will want to be in the roll-call scenario, since being able to see previous votes increases their chances of voting correctly, but as the group becomes more connected its performance will degrade due to the increased chance of information cascades, meaning that everyone becomes worse off. This dilemma generalises: for models with intermediate levels of connection, the more connected the groups, the more reliable individuals are, and the less reliable the collective position is. Zollman (forthcoming) establishes these results with a mathematical model, but given the empirical evidence for information cascades,12 and jury-theorem-type results13, we ought to take this kind of social dilemma seriously.

2.4.2. The Problem of the Second-Best

Let’s now turn to the relation between institutional epistemic properties and epistemic justice. Conflict emerges between these two goals because the ideally best institutional arrangements may not be the best response in non-ideal conditions, and may actually exacerbate injustice. This is a cousin of the problem of the second-best in economics: in non-ideal political conditions, trying to approximate the epistemically ideal institutional structures may lead to suboptimal outcomes.

Given our focus on social media, it will be helpful to see how the problem of the second-best shows up for the Habermasian ideal of the public sphere.14 One of the aims of Habermas’ *The Structural Transformation of the Public Sphere* (Habermas 1989) is to argue that the epistemic and political ideal for democratic society is a unified public sphere, in which citizens are able to together address issues of common concern on an equal footing, making appeals to public reason. Habermas is not so naïve as to think that any actually existing democracy has realised this ideal, but he does think that the discursive arenas of the eighteenth century—salons, coffee houses, and drinking halls—aspired to it, despite their actual exclusions. If one is attracted to the idea that the epistemic ideal for a democratic society is a unified

---

13 (Miller 1996).
14 In the interests of space, I will bracket the question of whether social media is best thought of as a proper part of the public sphere. My suspicion is that some parts of social media sites are clearly in the public sphere (politics discussions on twitter, sharing news articles on Facebook), others fall clearly into the private sphere (gossip about friends, private counterpublic groups), and many either fall hazily in-between, or shift abruptly from the private to the public. To say that a discursive space is part of the private sphere is not to say that it is irrelevant to the public sphere: Habermas’ discussion of private literary culture in the 18th century (Habermas 1989: 43-56)) makes clear that he takes the public sphere to rest on effective private spheres.
discursive space for airing common issues, then it would be natural to think that more unification in public discourse is a good thing, and more fragmentation is a bad thing. The idea of a unified public sphere has been an important ideal for social media and has been leveraged by many authors to argue that information fragmentation on social media is a bad thing.

The problem of the second-best emerges in this case because in non-ideal circumstances, more unification can actually undermine the epistemic workings of the public sphere. This line of argument is pushed by Nancy Fraser (Fraser 1990). Fraser points out that in non-ideal circumstances, women and minority groups face a number of connected barriers to participants, besides simply being excluded from discursive spaces, including restrictive conceptualisations of the bounds of ‘the political’, male-coded norms of political speech, and conceptual lacunae. Fraser argues that for actually existing societies characterised by political inequality, greater unification in the public sphere would compound these political problems. Building on the work of feminist historians and political theorists, Fraser proposes that the cause of political equality and a well-functioning public sphere would be better promoted by a network of counterpublic spaces, which allow minority groups to bracket the epistemic oppression of the principal public sphere, and develop the epistemic resources to both adequately represent their interests and contest the restrictions of the public sphere. Fraser’s critique of Habermas sets up a neat version of the problem of the second-best: Habermas proposes an attractive ideal, but moving from in that direction from our current non-ideal political system leads to bad consequences (and may actually undermine the realisation of the ideal). Given our current political situation, we would be better off focusing on an ideal which ameliorates our current problems, which Fraser argues is a network of counterpublic spaces.

2.4.3. Individualism and Epistemic Justice

The conflict between individualism and epistemic justice emerges because of individualism’s consequentialist structure. Veritism, for example, evaluates social systems by considering the average changes in accuracy for participants in that system, without regard to how those changes are distributed, or how they are brought about. Just as consequentialist accounts of moral goodness struggle with the distribution of value, and side-constraints on the pursuit of value, individualism about epistemic value struggles with the importance of equitable epistemic outcomes, and respect for epistemic agency. Let’s take these two ideas in turn.

---

15 Here Fraser anticipates several of the central themes of Fricker (2007) (See also Fricker 1998, 1999). Of course, work on epistemic oppression in Black Feminism antedates Fraser by at least a hundred years and has been ignored by white feminists (Dotson 2012, Berenstain 2020). On Fraser’s failure to deal with black counterpublics and for a more detailed typology of subaltern public spheres, see (Squires 2006).
16 See also (Collins 2000: 111-23) on the epistemic functions of safe spaces.
17 The third part of epistemic justice—inverted epistemologies—is easier to accommodate in a veritist framework, see (Mills 2007).
First, equality in epistemic outcomes (Fallis 2003, Coady 2010, 2017). Consider two systems: the first of which averages large increases to accuracy while decreasing the accuracy of a minority, and the second which yields positive changes in accuracy to all participants but with a lower average accuracy. Veritism predicts that the first system is better in all relevant respects, whereas we want to say that there is at least something bad about the more unequal system, and that at in at least some cases a more equitable epistemic system is more important than greater average accuracy. What goes for differences in outcomes between individuals also goes for differences in outcomes for different questions: epistemic systems can go wrong by being less reliable for questions about some social groups than others.18

Second, side-constraints to the pursuit of epistemic value. Much as moral rights against certain kinds of harm constrain the pursuit of good consequences, the right to epistemic autonomy constrains the pursuit of good epistemic consequences. If one epistemic system produces higher average accuracy than another, but by undermining the epistemic agency of minority groups, there is something bad about that system, and in at least some cases greater accuracy is trumped by concern for epistemic agency. Think of cases in which the accuracy-generating features of a system relies on excluding minorities, or on treating members of minority groups as epistemic instruments (rather than as inquirers).

3. The Social Epistemology of Social Media

According to critical social epistemology, evaluating social systems from an epistemic point of view involves the consideration of individual epistemic outcomes, institutional properties, and social justice, and in important sets of cases, the pursuit of these goals will lead to conflict (see fig 1).

---

18 This is one part of what goes wrong with cases of algorithmic bias (Fazelpour and Danks 2020). The ethics of question-choice is an important theme in feminist epistemology of science, see (Anderson 1995).
With this picture of the epistemic aims of social institutions in place, we can consider how it helps to illuminate debates about the epistemology of social media. I will do this in two stages: first showing how discussions of the epistemology of social media appeal to these different epistemic aims (section 3), before demonstrating how the tensions between these aims show up for social media (section 4).

3. Individualism, Institutionalism, and Epistemic Justice

First, we will consider some debates about the design of social media that demonstrate that concerns about individual epistemic goods, collective epistemic virtue, and structural epistemic justice play a role in debates about how to design and reform social media. These examples are picked because they are neat examples of appeals to different kinds of values, and because they help to illustrate the conflicts between different epistemic aims, not because these values are the most important for social media.

3.1 Individualism

Many changes to the rules and functionality of social media sites make appeals to individual epistemic goods (see Silverman 2019, Rosen and Lyons 2019). We will consider two examples: Facebook’s changes to the newsfeed in April 2019 to reduce the amount of ‘false news’, and Twitter’s changes to the retweet function in October 2019.19

In a press release from April 2019, Facebook announced a slew of changes to the platform that were designed “to fight false news”. These changes included: exploring the possibility of crowdsourced fact-checking, expanding the role of the Associated Press in fact-checking stories, reducing the reach of Facebook Groups that repeatedly share false stories, and incorporating a ‘Click-Gap’ signal into the newsfeed algorithm which promised to reduce the reach of domains which are more linked on Facebook than they are on the rest of the Internet. None of these interventions involve straightforwardly removing false content, but—setting to one side blustering rhetoric about ‘fighting’ misinformation—the substance of this proposal is to implement changes to the newsfeed algorithm to reduce the amount of false claims that show up on the newsfeed, with the proximal aim of ensuring that people who use Facebook come away with fewer false beliefs. Here Facebook is deploying a kind of error-avoiding veritism according to which social media platforms are better when they make their users beliefs less inaccurate.

19 These interventions involves a degree of epistemic paternalism, raising difficult questions about social media sites’ right to interfere with their users epistemic states (Castra, Pham, and Rubel 2020)
Social media companies are interested in other individual epistemic outcomes besides the accuracy. In October 2019, as part of its preparations for the US election, Twitter announced changes to the way users could repost other posts. At this point, Twitter users could click the retweet button and choose to retweet without comment, or quote tweet with a prefatory message. To encourage more ‘thoughtful amplification’, the retweet function was changed so that pressing the retweet button led straight to the quote tweet screen (although it was still possible to retweet without a comment by leaving the text box empty). The hope was that by adding friction to retweeting\(^{20}\) would encourage users to add their own commentary, “increase[ing] the likelihood that people add their “own thoughts, reactions and perspectives to the conversation.”\(^{21}\) This change was reversed in December 2020. In a blogpost, Twitter employees report that “we observed that prompting Quote Tweets didn’t appear to increase context: 45% of additional Quote Tweets included just a single word and 70% contained less than 25 characters.”\(^{22}\) We can see this change as motivated by an interest in users making more justified contributions to conversation. By adding friction into the retweet function, Twitter users were discouraged from quick and unjustified retweets, and were encouraged to take a second to think about retweeting by adding supporting (or rebutting) reasons to the claim that they were retweeting.\(^{23}\) This example highlights that individualists can be interested in other epistemic goods besides accuracy, such as whether beliefs are justified, and the quality of contributions to collective inquiry.

### 3.2 Institutionalism

Alongside concerns for individual epistemic goods, discussions of social media demonstrate concerns with institutional epistemic properties. We will focus on two particularly clear examples: Cass Sunstein’s view that the ideal for social media is a unified public sphere, and Facebook’s appeal to the value of connectedness.\(^{24}\)

In a series of remixes of his book *Republic.com* (Sunstein 2001, 2007, 2017), Cass Sunstein raises worries about the ways in which information technologies for ‘filtering’ information create a fractured public sphere that is unable to play its role in deliberative democracy (see also Pariser 2011). He has something of a shifting target: in *Republic.com* his focus was on personalised news and niche websites, *Republic.com 2.0* changed focus to blogs, and #Republic was squarely focused on social media.

---

\(^{20}\) On the importance of epistemic friction to social media, see (Ashton 2020, forthcoming).

\(^{21}\) (Gadde and Beykpour 2020a)

\(^{22}\) (Gadde and Beykpour 2020b)

\(^{23}\) Although it is tempting to assimilate retweeting to assertion (Rini 2017), retweeting appears to be a distinct speech act that only sometimes involves indirect endorsement of the embedded claim (Marsili 2019), (Pepp, Michaelson, Sterken forthcoming).

\(^{24}\) For a book that applies tools from network epistemology to problems around misinformation, see (O’Connor and Weatherall 2019).
In Sunstein’s view, a well-functioning deliberative democracy requires certain supporting institutions: general-interest venues for disseminating information that establish a common ground of accepted beliefs and enable the serendipitous discovery of novel views. His concern is that information technologies that allow for personalised filtering herald the death of these institutions (he vacillates between concerns with current technologies, and worries about future developments). The worry is that a media system organised around consumer interests will focus people’s attention on special-interest forums, undermining the common ground through the filtering of information through search engines, hashtag-based social media, and personalised newsfeeds, ultimately replacing the unified public sphere with partisan news sources, and individualised information diets. Sunstein argues that this fragmentation of our epistemic landscape leads to a number of problems, including echo chambers, group polarisation, persistent false beliefs, cybercascades, a lack of shared experiences, and violent extremism. He canvasses a number of policy proposals for the internet including websites designed for civil deliberation, subsidies for public internet websites, encouraging viewpoint diversity on partisan websites, and ‘opposing viewpoint’ and ‘serendipity’ buttons for social media sites (Sunstein 2017: 213-233).

Sunstein situates his work in a Republican tradition of American legal scholarship that grounds the value of free expression in the epistemic value of public discourse, but there is a strongly Habermasian flavour to Sunstein’s critique ((Sunstein 2017: 46-7, 153-5) see also (Habermas 2006: fn3)). Like Habermas, Sunstein claims that the ideal for a deliberative democracy is a unified public sphere in which all citizens participate on a relatively equal footing, have shared experiences that support a common ground of beliefs, and have chance encounters with a wide range of fellow citizens with different views. The problem with special-interest websites, blogs, and social media is that they lead to a fractured public sphere that undermines collective democratic problem-solving. Sunstein’s critique of the contemporary internet is robustly institutionalist. He uses metaphors about public space to frame his argument, and his central question is “what are the social preconditions of a well-functioning system of democratic deliberation?” (Sunstein 2017: 5).

Institutional epistemic features also play a role in how social media companies present themselves to the public. Since at least 2010 Facebook has presented itself as promoting the social purpose of connecting users (van Dijck 2013). Here’s Mark Zuckerberg’s celebrating Facebook reaching two billion users in 2015:

I’m so proud of our community for the progress we’ve made. Our community stands for giving every person a voice, for promoting understanding and for including everyone in the opportunities of our modern world. A more open and connected world is a better world. It brings stronger relationships with

---

25 In other work, Sunstein has criticized Habermas, arguing that prediction markets are preferable to democratic deliberation (Sunstein 2006).
those you love, a stronger economy with more opportunities, and a stronger society that reflects all of our values.26

In this passage, Zuckerberg highlights the political and epistemic benefits of connectivity, for individuals and societies. Over the years, Facebook’s use of ‘connectivity’ has acquired an expansive set of associations, referring to how many people are using Facebook, projects to provide free internet access, and a sense of social closeness. In a post from 2017 entitled *Bringing the World Closer Together*, Zuckerberg ties these associations together through the history of the site:

I started Facebook to connect my college. I always thought one day someone would connect the whole world, but I never thought it would be us. I would have settled for connecting my whole dorm. We were just college kids. But we cared so much about this idea -- that all people want to connect. So we just kept pushing forward, day by day, just like you.27

It is no accident that Facebook is designed for users to become highly connected. José van Dijck argues that Facebook’s rhetoric of social connectedness provides ideological cover for an economic model which is based on extracting data from users’ social graphs (van Dijck 2013, see also Zuboff 2019). By designing a platform that incentivises users to connect with many others, along with features like the ‘like’ button which tracks users across the rest of the internet, Facebook is able to harvest more information about their users, which it can use to offer its customers highly targeted adverts.

Notwithstanding Facebook’s commercial interest in a highly connected userbase, this raises a substantive issue for epistemologists: is a highly connected group of people an epistemically virtuous social structure?28 Zuckerberg suggests that ‘highly connected’ groups promote understanding, and a stronger society, but as we shall see below (in section 4.1.), connectedness raises difficult issues about the relation between individual and collective epistemic value.

3.3. Epistemic Justice

In section 2.3., we distinguished three parts of epistemic justice: epistemic oppression, equality of epistemic goods, and inverted epistemologies,. In thinking about these issues, it is important to consider how the epistemic functioning of social media platforms interacts with the way technology reinforces oppressive relationships, and enact forms of racial profiling, making them complicit in enacting what Noble calls *technological redlining* (Noble 2018, see also O’Neill 2016, Eubanks 2018, Benjamin 2019, Mullaney, Peters, Hicks, Philip 2021).

---

26 (Zuckerberg 2015)
27 (Zuckerberg 2017)
28 There is also an open question about how connected Facebook actually is, given the algorithmic filtering of the timeline.
We can find many examples of social media sites abetting inverted epistemologies, both by permitting certain kinds of organising, and by providing the technological basis of white ignorance. As Safiya Noble points out in the epilogue of *Algorithms of Oppression* (Noble 2018), a significant part of the issue of misinformation—which is often ignored in discussions of ‘fake news’—concerns technologically-abetted ignorance about marginalised groups. If sites fail to moderate racist hate speech (as the majority of social media companies have), they are effectively providing an organising service for white supremacists who have a history of being ‘innovation opportunists’ (Daniels 2009, 2018). White supremacist organising on social media is not merely permitted, but tacitly abetted by algorithmic systems which appear to promote racialised ignorance. For example, it has been suggested that YouTube has a bias towards ‘extreme’ content, creating a situation in which the recommended videos for fairly banal videos will quickly lead users down a pipeline of increasingly white supremacist content (Tufekci 2018, Alfano, Carter, and Cheong 2018, Munn 2019). If social media sites are designed for algorithmic self-radicalisation, they are complicit in the inverted epistemology of white supremacy.  

We find no shortage of examples of marginalised groups on social media having their capacity to contribute to inquiry being undermined. Social media companies permit significant amounts of online harassment, which has disproportionate harms on women, especially women of colour, lesbian women, and trans women. Sarah Sobieraj argues that the online harassment of women should be understood as “patterned resistance to women’s public voice” (2018: 1701, see also 2020). She argues that this resistance employs common misogynistic strategies of intimidation, shaming, and discrediting that make gender and women’s bodies salient in online spaces.  

She argues that the cumulative effect of misogynistic harassment is to “silence women, undermine their contributions to digital discourse, press them out of valued digital publics, and create a climate of self-censorship that mirrors the calculations women make in physical public spaces about what is safe and what is risky.” (Sobieraj 2018: 1709-10). This means that social media platforms are sites of epistemic oppression, and suggest that a significant proportion of the responsibility accrues to the companies who designed these systems. Given the prevalence of technological redlining in other algorithmic systems, it is plausible that the algorithmic systems that sort timelines, suggest connections, and contribute to content moderation will further undermine the epistemic agency of marginalised groups. Although commercial secrecy means that we lack detailed information about these systems, it is telling that the automated systems for detecting hate speech produce high rates of false positives when used on African American Vernacular English (AAVE), due to biases in the training data (Sap et al 2019, Davidson, Bhattacharya, and Weber 2019, Kim et al 2020).

---

29 For an alternative take that stresses the role of social networks in radicalization, see (Lewis 2018).

30 Sobieraj focuses on the strategies of misogyny, but there will be further stories to be told about misogynoir and transmisogyny.
4. Conflicts between the Epistemic Aims of Social Media Platforms

Let’s now turn to the conflicts between the epistemic aims of social media, considering what light the independence thesis, the problem of the second-best, and the conflict between epistemic justice and maximising epistemic goods can shed on debates about social media.

4.1. The Independence thesis

In section 2.4.2, we saw that the independence thesis means that the methodological prescriptions for individuals and groups can come apart. While we shouldn’t mistake mathematical models for reality, but there are good reasons to think that currently existing social media sites exemplify this dilemma, in the sense that social media networks are good for individual epistemic outcomes, but bad for collective-level epistemic performance.\(^{31}\)

We saw in section 3.2, that Facebook presents itself as promoting connectedness amongst its users, encouraging users to create connections with many others. This kind of network structure does offer some epistemic benefits. In a highly connected network, individuals have access to a greater amount of information, can expect to hear of interesting news in a timely manner, and will be able to make use of coverage-supported reasoning (see Goldberg 2010 C6). However, the fact that highly connected networks provide epistemic goods for individuals does not mean that they are without problems at the institutional level. Highly connected networks can reach consensus too quickly (Zollman 2007, 2010), and can lose reliability due to informational cascades (Zollman forthcoming). Although this would take more evidence to back up, we might speculate that the reason why high-connection sites like Facebook and Twitter remain popular despite their well-publicised problems is that they trade off individual epistemic goods against our collective epistemic interest. We gain fast and reliable information about what our friends are up to and breaking news, but at the cost of information cascades and false but surprising stories spreading quickly across closely connected networks.\(^{32}\) People continue to use these sites not because they are irrational or addicted, but because they are trading their own epistemic interests off against the good of the collective.

4.2. The problem of the second-best

---

\(^{31}\) For another case in which individual and collective dimensions of evaluation diverge, see (Nguyen 2020).

\(^{32}\) We can also find examples of the reverse of this effect. The best candidates for sites that take advantage of the Condorcet Jury theorem are prediction markets and betting exchanges. While the outputs of prediction markets might be more reliable than the average user, these sites allow very restricted (or no) communication between users to ensure that bets are made as independently as possible, meaning that their collective performance comes at the cost of restricting individual-level evidence.
The conflict between designing social media for collective epistemic virtue and for structural epistemic justice shows up in debates about the online public sphere. We saw in section 3.2. that Sunstein suggests that the ideal for social media in a deliberative democracy is a unified public sphere in which citizens shared a common ground of accepted beliefs and are exposed to a wide range of perspectives. The unified public sphere might be a good ideal for a society that has already achieved conditions of structural epistemic justice. But, as Fraser (1990) argues, in which are characterised by endemic exclusion and silencing in the public sphere, a unified public sphere would only compound epistemic oppression.

Applying Fraser’s analysis to online discourse, we might suggest that for societies characterised by epistemic oppression, the proper ideal is not a unified public sphere realised by one social media site, but rather a network of counterpublic spheres representing an overlapping set of groups’ interests. Although unification might be the epistemic ideal for a just society, fragmentation will promote epistemic justice in an unequal society like ours. Work in media studies has been heavily influenced by the public sphere tradition, and there is a rich body of work on online counterpublics, including curated blocklists on twitter (Geiger 2016), Black twitter (Graham and Smith 2016, Brock 2020 C3), hashtag networks (Jackson, Bailey, Foucault Welles 2020), blogs (Steele 2016a, 2016b, Gabriel 2015), and BlackPlanet (Byrne 2007). Some of this literature has a normative bent, arguing that Online counterpublic spaces can play a role in resisting epistemic oppression by providing safe havens and support networks (Geiger 2016, Sobieraj 2018), allowing marginalised groups to contest media narratives (Steele 2016a, 2016b, Gabriel 2015), and allowing marginalised groups to organise (Jackson, Bailey, Foucault Welles 2020). We can think of this literature as providing us with a first pass at an account of what effective online counterpublics look like.

4.3. Individualism and Epistemic Justice

In section 2.4.3., we saw that the consequentialist structure of individualist approaches to social epistemology means that individualism is unable to account for the importance of equality of epistemic outcomes and certain kinds of epistemic rights. Conflicts between majority epistemic interests and epistemic justice can in principle break either way, and there may be some cases in which the epistemic interests of the majority justify some undermining of epistemic agency. However, in the majority of realistic cases the badness of epistemic oppression and unequal distributions of epistemic outcomes trump maximising aggregate epistemic outcomes.34

33 To be fair, Sunstein does note the importance of deliberative enclaves to public discussions, especially in context of epistemic oppression (2017: 85-89, 254). However, his policy proposals focus on promoting a unified public sphere, and neglect the question of how to promote well-functioning counterpublics.

34 An analogy: public buildings which are inaccessible for people who use wheelchairs or other mobility aids are a bad thing from the point of view of inclusion and equal access. In the majority of cases ensuring equitable access comes at a reasonable cost, and does not
Social media companies give everyone who uses them access to the same basic services in the form of the public website, but these services may work differently, or be differentially accessible for different groups of people. Many social media sites are designed primarily for people who speak English, meaning that they offer importantly different services to non-English speakers. These differences have epistemic consequences: if anti-misinformation efforts are focused on posts in English, then users who do not speak English or primarily speak another language will not benefit from these efforts. Many social media sites are also poorly designed for users with sensory impairments and other accessibility issues. For example, images on Twitter do not have automatic descriptions, and YouTube’s automated captions have frequent errors and fail to distinguish between different speakers. Issues about the distribution of epistemic goods should be situated as part of a broader concern with the differential benefits and burdens of social media sites, as part of a wider concern with design justice (Constanza-Chock 2020).

Besides concerns with the distribution of epistemic goods between individuals, we might also be concerned with the distribution of epistemic goods regarding different questions. Algorithmic newsfeeds typically promote certain kinds of content—posts from news sources, posts from friends, posts from smaller groups—and by so doing, they will prioritise users having accurate beliefs about different kinds of questions. Facebook’s reorganisation of their newsfeed in 2018 prioritised posts from connected users over posts from public sources including news media and businesses. One of the effects of this change is that using Facebook’s newsfeed will have a greater impact on users’ beliefs about what their friends are up to than their beliefs about current events. While it is possible that this change might be a good thing from the point of view of average accuracy (assuming that content about one’s friends is more accurate that news stories), we might be concerned that this change means that Facebook fails to provide accurate information relative to the set of important questions we collectively care about, which includes political questions.

Finally, we ought also to be concerned about the ways in which pursuing majoritarian epistemic interests can indirectly lead to the epistemic exclusion of marginalised groups. Epistemic accountability practices often lead to good epistemic outcomes by allocating non-epistemic costs to people who act in ways that aren’t socially beneficial (see Frost-Arnold 2014). For example, we might promote the decrease the ability of people who do not use wheelchairs from using public buildings. However, in some special cases the benefits of ensuring accessibility for minority groups decreases the overall accessibility of a building. If the belltower of a gothic cathedral can only be accessed by a steep circular staircase, making the tower accessible for wheelchairs (for example by replacing the stair with a lift) might mean that fewer people are able to access the tower.

35 (Mosseri 2018)
36 This paragraph is heavily influenced by Karen Frost-Arnold’s discussion of the epistemic benefits of anonymity (Frost-Arnold 2014). However, whereas Frost-Arnold articulates the harm of excluding marginalized groups in veristic terms, I focus on understanding this harm in terms of epistemic exclusion.
goal of people sharing accurate information by allocating social sanctions to people who are discovered to have shared false information. Epistemic accountability practices are extremely common online: we might think of call-outs that aim to embarrass a unreliable user, calls to unfollow persistently deceptive users, as well as actions taken by sites to restrict users who share false information. While well-designed accountability practices may be good from the point of view of maximising epistemic goods, they can have pernicious effects on marginalised users. These users may be less able to bear the costs of accountability practices, may receive disproportionate negative effects for flouting shared epistemic norms, and may be more likely to be incorrectly accused of flouting epistemic norms. This unequal distribution of costs can have a number of negative effects: marginalised users may change what they say (or stay silent) to avoid the costs of perceived norm flouting, avoid contentious topics, or simply move to different social media sites without such rigorous accountability practices. Self-exclusion from spaces of public discourse and self-silencing owing to accountability practices are doubly problematic: they undermine users’ epistemic agency by making discursive spaces inhospitable environments, and it removes the specialised knowledge of marginalised people, meaning that epistemic communities become less diverse (Frost-Arnold 2014). While I don’t want to suggest that epistemic exclusion always trumps the pursuit of maximising good epistemic outcomes, but acknowledging the costs of practices that are good for the majority of users complicates the assessment of social practices, highlighting the importance of centring marginalised groups in systems-oriented social epistemology.

5. How to Design an (Epistemically) better Social Media

Critics of social media who are in the grip of an exceptionalist attitude often structure their critique around a specific epistemic vice of social media (the large amount of false posts, the existence of filter bubbles, algorithmic boosting of certain kinds of content), which their amelioratory proposals fixes by making online environments resemble our offline epistemic practices (adding reputation systems, ‘mixing up’ filter bubbles by adding random mixing of users, prioritising human editing over algorithmic filtering). While some of these proposals may be beneficial, our discussion of the conflicts between the different epistemic aims of social media highlights the importance of thinking in a joined-up way about the different parts of the epistemic evaluation of social media. Adding a reputation system to a social media might help to reduce the amount of false information posted (Rini 2017), but it might also reduce the amount of true information posted, have chilling effects on the speech of marginalised groups, and undermine epistemic diversity (Frost-Arnold 2014). Mixing up information may break up problematic epistemic bubbles (in the sense of Nguyen 2020), but it also runs the risk of undermining the formation of well-functioning counterpublics that rely on filtering out certain kinds of posts (see

---

37 This phenomenon is closely connected to what Dotson calls testimonial smothering, which occurs when speakers truncate their testimony due to their audience’s testimonial (in)competence (Dotson 2011).
Prioritising human editorial work over algorithmic filtering might help to ameliorate the algorithmic boosting of certain kinds of content, but it runs the risk of replacing it with human-driven inverted epistemologies (see Frost-Arnold forthcoming). The epistemic evaluation of social systems is complex, and often involves trading off different epistemic goods against one another. One lesson from science and technology studies is that technology is never neutral, and often embodies political values (Winner 1980, Huebner and Liao 2021), and that is as true in the domain of epistemic evaluation as in any other.

How we ought to resolve conflicts between the different epistemic desiderata of social systems is a difficult question. In the absence of a compelling meta-epistemological theory about the fundamental values are in the domain of epistemology, we have a couple of options. We might think that conflict cases are irresolvable dilemmas, or we might appeal to values from outwith epistemology to resolve the conflict cases. I think that our discussion of the social epistemology of social media, and an appreciation of the political valences of technology points towards the second option. The conflict cases we have considered centre around political and ethical conflicts in which the needs and interests of individuals, different groups, and the ethical value of knowledge-producing institutions need to be weighed up against one another. This is one reason why work in critical technology studies is vital to the social epistemology of social media (and the epistemology of technology more broadly). This work both highlights the ways in which technological systems are complicit in epistemic oppression and develop accounts of the conflict between epistemic justice and other epistemic values. For example, the point of critical work on algorithmic systems is not to argue that they are always unreliable—although they are often less reliable that their creators would like us to think—but to contend that their general reliability comes at the cost of ethical and epistemic harms to minority groups and hides specific kinds of reliability on questions about minority groups (Noble 2018).

This paper aimed to set out the groundwork for thinking about the social epistemology of social media, and in so doing to provide some case studies in anti-exceptionalist thinking about the epistemology of social media. There is much more work to be done exploring different aspects of social media, drawing on literature from other critical projects (for more work in this theme, see Frost-Arnold forthcoming). However, we are hopefully in a better position to see what kind of thing a critical social epistemology of social media might be. The critical social epistemology of social media is an ethical project because the design of social media systems involves conflicts between different epistemic values. It is a conflictual and political project that involves taking decisions about the priority of different groups’ interests, and the interests of minority groups ought to be centred in this inquiry. It is a pluralistic project that needs to draw on different methodologies from within social epistemology, and an interdisciplinary project which will draw ideas and results from a range of related disciplines. It is also a project of imagination that needs to break out of the boundaries of what currently existing social media sites
look like to imagine how they might be otherwise.\textsuperscript{38} Situating the evaluation of social media sites in the general evaluation of epistemic systems hopefully goes some way towards opening up our imaginations.\textsuperscript{39}

References


Ashton, Natalie Alana (2020 Why Twitter is (Epistemically) better than Facebook, \textit{Logically.ai} (26\textsuperscript{th} Nov 2020) https://www.logically.ai/articles/why-twitter-is-epistemically-better-than-facebook

Ashton, Natalie Alana (MS) Productive Online Environments.


\textsuperscript{38} On the political work of imagination, see (Rorty 1983, Stoetzler and Yuval-Davis 2002, Biss 2013, Dotson 2018).


Habermas, Jürgen (1989) The Structural Transformation of the Public Sphere, Thomas Burger and Frederick Lawrence (trans), Polity Press, Cambridge

Jackson, Sarah J., Bailey, Moya, and Foucault Welles, Brooke (2020) #Hashtag Activism: Networks of Race and Gender Justice, MIT press, Cambridge MA


McKenna, Robin (MS) Non-Ideal Epistemology


doi: http://dx.doi.org/10.5210/fm.v24i6.10108


Mosseri, Adam (2018) Bringing People Closer Together, *facebook Newsroom*

Mullaney, Thomas, Peters, Benjamin, Hicks, Mar, Philip, Kativa (2021) *Your Computer is on Fire*, MIT press, Cambridge MA


Pepp, Jessica ; Michaelson, Eliot & Sterken, Rachel (forthcoming). On Retweeting


Steele, Catherine Knight (2016b) Signifyin’, Bitching, and Blogging: Black Women and Resistance Discourse Online, in *The Intersectional Internet: Race, Sex, Class, and Culture Online*, Safiya Umoja Noble and Brendesha M Tynes (eds) Peter Lang New York pp. 73-94


Technology, & Public Life (CITAP), University of North Carolina at Chapel Hill. http://citap.unc.edu/critical-disinfo


Winner, L. (1980). Do artifacts have politics?. *Daedalus, 121*-136.


Zollman, Kevin J. S. (forthcoming) *Network Epistemology*

Zuckerman, Ethan and Rajendra-Nicolucci, Chand (2020) Beyond Facebook Logic: Help us map Alternative Social Media! Knight First Amendment Institute, Columbia University Blog https://knightcolumbia.org/content/beyond-facebook-logic-help-us-map-alternative-social-media