Epistemic communities and trust in digital contexts*

Comunidades epistémicas y confianza en contextos digitales

ANTONIO GAITÁN TORRES**

Resumen: Este comentario se centra en la noción de 'comunidad epistémica' y en su rol para sustentar el argumento general que Karen Frost-Arnold presenta en *Who Should You Be Online?* (OUP, 2023). En la primera sección se presenta el argumento general de WSYBO, esbozando la estructura general del libro y sus conceptos centrales. En la segunda sección se distinguen tres sentidos posibles de 'comunidad epistémica' – sistémico, agregado y grupal. En la tercera sección se exploran las tensiones en torno al sentido grupal de comunidad epistémica. Se atenderá a dos formas en las que ciertas comunidades epistémicas cerradas y organizadas en torno a una identidad compartida pueden desviarse de los ideales epistémicos que guían el proyecto de Frost-Arnold. Algunas comunidades cerradas pueden organizarse en torno a dinámicas excluyentes. Otras comunidades pueden organizarse en torno a debates o controversias, afectando a la

Abstract: This commentary focuses on the notion of 'epistemic community' and its role in underpinning the general argument that Karen Frost-Arnold presents in *Who Should You Be Online?* (OUP, 2023). The first section presents the general argument of WSYBO, outlining the general structure of the book and its central concepts. In the second section, three possible senses of 'epistemic community' are distinguished – systemic, aggregate and group-oriented. The third section explores tensions around a variety of group-oriented epistemic community. It will address two ways in which certain closed epistemic communities organized around a shared identity can deviate from the epistemic ideals that guide Frost-Arnold's project. Some enclosed epistemic communities can be organized around exclusionary dynamics. Other enclosed epistemic communities may organize and grant membership around debates or controversies, affecting

** Universidad Carlos III de Madrid, Profesor Titular, Departamento de Humanidades – Filosofía, Lengua y Literatura, Facultad de Humanidades, Comunicación y Documentación. Sus líneas de investigación incluyen la ética normativa y la meta-ética, la filosofía experimental y la epistemología social y política. Últimas publicaciones: Gaitán, A. (2024). "La psicología de las emociones políticas', en Gómez-Ramos, A. Velasco, G. (eds.). *Atlas de Emociones Políticas,* Madrid Trotta; Gaitán, A. Viciana, H. Aguiar, F. (2023). "The Experimental Turn in Moral and Political Philosophy", en Viciana, H. Gaitán, A. Aguiar, F. (eds.). *Experiments in Moral and Political Philosophy*, Londres, Routledge. Correo electrónico: agaitan@hum.uc3m.es

calidad deliberativa de esos debates. El potencial epistémico de las comunidades epistémicas cerradas y organizadas en torno a una identidad compartida depende en gran medida de evitar estas dos desviaciones.

Palabras clave: Confianza, comunidades epistémicas, epistemología social, epistemología orientada hacia sistemas.

the deliberative quality of those debates. The epistemic potential of closed epistemic communities organized around a shared identity depends largely on avoiding these two deviations.
Keywords: Trust, epistemic communities, social epistemology, system-oriented epistemology.

In *Who Should You Be Online* (OUP, 2023 – WSYBO hereafter), Karen Frost-Arnold proposes 'a socially situated epistemology for the internet' (WSYBO, p. 203); 'a philosophical, activist and non-ideal framework' (WSYBO, p. 5) that should contribute to improve our interactions in various digital contexts. To articulate this framework, Frost-Arnold appeals to several philosophical concepts, as well as a huge wealth of empirical evidence focused on our behavior and attitudes in digital contexts. WSYBO is full of concrete examples, always illustrating far-reaching philosophical claims in an entertaining and accessible way. All this is achieved in a constant dialogue with disciplines such as Anthropology, Sociology, Political Science, Communication or Digital Studies.

There are many interesting aspects of WSYBO that could be commented on, either on some long-standing debates in Epistemology (trust, epistemic autonomy, epistemic injustice), or on more novel topics and lines of research (the epistemic risks of moderation, the group biased nature of fake-news, the dangers and possibilities of imposture in digital contexts, 'lurking' as an epistemic practice and a long etcetera). Here I will focus on the notion of 'epistemic community', a concept that I find especially interesting when it comes to understanding and evaluating the project that Frost-Arnold develops in WSYBO.

1. The project

The central thesis of WSYBO is that any 'epistemic system' (press, education, social media, etc.) that seeks to promote the generation of objective knowledge – free of biases and beyond partial perspectives – must be capable of implementing two great ideals. First, it must be diverse in a strong sense, that is, the epistemic system must prioritize not only a plurality of perspectives; It must also include the perspective of the disadvantaged or oppressed groups. Secondly, and as a necessary condition for getting the epistemic benefits to be brought by this strong diversity, it must promote and protect networks of trust that ensure the expression of subaltern or oppressed voices (WSYBO. p. 75).

It is convenient to stop at this second condition. According to Frost-Arnold, diversity in the strong sense has no real effects for the generation of objective knowledge if oppressed groups are not part of the trust networks established between different epistemic communities. For the inclusion of oppressed voices to be effective, trust networks must be robust in three different senses. Oppressed groups must trust their own perspective, have the trust of other members of the community, and themselves trust those practices that the community establishes to encourage the debate and the inclusion of different voices and perspectives (WSYBO. p. 76).

Assuming this general framework, much of WSYBO can be read as a catalog of characters or ideal types that would undermine networks of trust at one of its three vertices. Chapter two, for instance, explores a variety of 'imposture' in digital contexts in which an oppressed perspective is supplanted with the goal of fostering stereotypes about such perspective. During the Arab uprisings of 2010, a blog attracted the attention of many Western journalists and political commentators. The title was very suggestive: *A gay girl in Damascus* told the story of the Syrian revolution from the perspective of Amina Arraf, a lesbian activists of Syrian-American origin. Amina narrated in first person her experiences during the uprisings, giving a first-hand account of the ongoing political insurrection. Months later it would be discovered that the blog was actually a fake and that its author, Tom MacMaster, was an American student living in Scotland. MacMaster's imposture, regardless of his real motivations, surely eroded the confidence of many of the readers of the numerous political activists blogs that reported on the ground and in real time about what was happening in Syria. And this impoverished the understanding of the Syrian political uprising during those years.

WSYBO caters in several characters like this, always asking about their effects on those networks of trust necessary for the production and dissemination of objective knowledge. Attention is also paid to the role that the architecture and affordances of digital environments would play in protecting or eroding these networks of trust. The possibility of interacting in spaces protected by anonymity, or the possibility of controlling who participates in those forums in which oppressed groups are discussing problems related their oppression, allows such minority groups to safely develop critical and emancipatory perspectives. The last chapter, dedicated to lurkers, offers an example of these contexts and dynamics.

Frost-Arnold's project is carried out by taking for granted several concepts and debates. Within this territory of implicit assumptions, there is a word that appears a lot of times in WSYBO. I am referring to the notion of 'epistemic community'. Almost all the actors who appear in this book are not interacting in the vacuum. By contrary, they are inserted, more or less formally, into different epistemic communities. But what is an epistemic community? And how are epistemic communities fitting into the general project outlined above? Do the trust networks that enable the generation of objective knowledge depend on a certain variety of epistemic community? And if so, how to understand that dependency? I am sure the author can answer all of these questions, but even so let me develop here some tentative criticisms for the sake of having a better picture on the meaning and varieties of epistemic communities in WSYBO.

2. Three senses of epistemic communities

Three broad senses of 'epistemic community' can be traced throughout WSYBO. According to the first, *a systemic or structural sense*, an epistemic community is equivalent to a socio-epistemic system of rules, conventions, techniques, affordances and practices that are organized (at least partially) for the promotion of knowledge. Facebook and X are epistemic communities in this structural sense. They stand for digital spaces

designed and organized, at least in part, with the aim of sharing information and promoting debate - two goals that are essential for promoting objective knowledge in Frost-Arnold' sense[1].

In a second sense, epistemic communities are sometimes referred in WSYBO as *aggregates of opinions around a particular debate or topic*. Some threads on X in which disagreements develop over some public controversy offer a clear example of an epistemic community in this second sense. In these aggregates of opinions, sometimes very numerous, interaction is 'more or less' limited to the exchange of points of view on a given topic, for example, police brutality (WSYBO. p. 60). These epistemic communities can persist over time, forming stable spaces in which we can find and supply information on a certain topic.

Finally, in WSYBO a third sense can be traced, according to which an epistemic community refers to *a group organized around a topic, debate, controversy or shared social identity*. This is the sense that I am interested in exploring in this commentary. Now we are not referring to a mere aggregate of opinions, as in the previous case, but rather to spaces or communicative structures in which a *group* of people interact on a regular basis and in which we find different degrees of organization. The organization of these groups can be informal or regulated, allowing for different levels of openness and inclusion (WSYBO. p. 161). The stability of the interactions between members of an epistemic community can also vary, moving form longstanding groups to more momentaneous assemblies. And importantly, epistemic communities in this 'groupish' sense can be organized around a topic or interest, but also around a shared social identity.

A blog about the Second World War in which a significant number of users participate regularly constitutes an epistemic community in this third sense, that is, a group organized around a specific *topic*. The limits between the second sense seen above and the one we are now considering are sometimes blurry, but certain features can give us a clue as to when we are dealing with an epistemic community organized as a group and when we are facing an aggregate of opinions around an issue. The administrator of the blog we are considering could, for example, control who participates or instead have a more open policy, allowing to anyone who is interested in the topic to post opinions in the blog. To the extent that the organization of a blog allows the administrator to control who post on the topic, a blog is clearly different from a thread of X on the same topic. Frost-Arnold writes:

> "(…) blogs can create smaller epistemic communities by hosting active comment threads. Commentators often engage in a critical dialogue with one another, another potential objectivity enhancing practice". (WSYBO. p. 80)

In other epistemic communities, by contrary, the criteria for belonging do not have so much to do with specific debates or topics as with *explicit affiliation to a shared identity*. A group of black women who start a blog to share their experiences of discrimination in

---

1   The interesting discussion around moderation contained in the second chapter refers to this structural sense of epistemic community (p. 32). The epistemic risks of moderation (the silencing of minority voices, for example) also naturally refer to this structural or systemic level. The numerous references to inequalities between groups or the underrepresentation of a certain community within a social network are also examples of this first sense of epistemic community (p. 34)

the workplace and who are reluctant to allow the participation of non-black women offers a clear example of this type of epistemic communities (WSYBO. p. 170).

Communities organized around a shared social identity appear in various parts of WSYBO. The assumption is that, at least when they are not organized in an excessively closed or opaque way (allowing some non-black women to participate, or facilitating the 'lurking' of their interactions by external observers), they enact spaces that would facilitate the generation of objective knowledge (WSYBO. p. 166). In the case we are currently focusing on, the epistemic community of black women would offer first-hand information about the living conditions and experiences of an oppressed group. Lurkers could access this information and transmit it to other epistemic communities, enriching the overall debate (in the systemic sense noted above) with information about the group's effective situation of exclusion, or about their own perception of that situation. This information may be essential to guide and modulate possible policies aimed at mitigating or correcting such situation. Even if interactions with other epistemic communities are somehow reduced, these closed epistemic communities organized around identity-markers can generate long-term epistemic benefits. By offering security to its members and a basic feeling of understanding, they allow that certain experiences can be articulated more precisely.

These three varieties of epistemic communities appear in WSYBO. And the central argument outlined above can, in fact, be refined by considering these three broad varieties. The generation of objective knowledge evaluated in a systemic way points to the systemic sense outlined above. Diversity of opinions, in the strong sense that concerns Frost-Arnold, can be exemplified in specific contexts in which an issue is debated (second sense of epistemic communities), but it can also refer to the composition of groups with different degrees of opening (third sense of epistemic community). And the same can be said about trust, the oil that allows that diversity can have real effects on the promotion of objectivity. Sometimes, trust has to do with the ability to intervene in open deliberative spaces without suffering verbal attacks or threats (second sense). Other times, however, trust requires the protection offered by more enclosed groups, which allow certain critical opinions to mature and be articulated effectively (third sense). As I noted above, one of the most valuable things of WSYBO is the way it precises how these enclosed groups could also be of critical importance for generating objective knowledge in the broader epistemic community (first sense).

At a more general level, the evolution of our perception of digital contexts, from a distant period of initial optimism to the current wave of pessimism (which Frost-Arnold dates to around 2016), can also be interpreted in light of the three senses of epistemic community outlined above. The initial promises undoubtedly pointed to the potential of digital forums, understood as open spaces of deliberation, to generate objective and open knowledge (first and second sense) (Sunstein 2009). The current pessimism around social media points, by contrary, to the proliferation of powerful groups interfering with the free exchange of perspectives and ideas (third sense) (Habgood-Coote 2024).

As I have suggested above, Frost-Arnold assumes part of this general narrative about our perception of the potentialities of digital spaces, but she does not endorse an additional move, common in many conservative analyses of the epistemic potential of the internet. Frost-Arnold does not believe that closed groups organized around identity markers are necessarily positing a danger to the generation of objective knowledge. Under a systemic

perspective they can promote objective knowledge. This precise claim, framed in epistemic terms, is one of her major contributions in WSYBO.

In what follows I will focus on some epistemic communities organized around a shared identity. Although Frost-Arnold is correct when she accentuates the epistemic advantages of some of these communities, her discussion blurs some risks associated with other possible instances of identity-based epistemic communities. I will argue that keeping these risks in sight could help us to better understand the potential of Frost-Arnold's general argument. The goal of the following section, in any case, is mostly exploratory. I am sure the author can deal with most of the criticisms I'll unfold.

3. Identity-based epistemic communities - the good, the bad, and the ugly

There is a general and straightforward way to articulate what concerns me in relation to the 'groupish' and identity-related sense of epistemic community. This strategy would point out, in a nutshell, that our group tendencies involve dispositions and beliefs that have bad epistemic consequences. Favoring the members of our group, limiting contact with members of other groups, openly discriminating against them when distributing resources, time and attention, are paradigmatic examples of epistemically deviant 'coalitional' dispositions (Boyer 2020). Although there are some senses in which some group-centered bias can have epistemic value (Rini 2017), coalitional dynamics have in principle little value in generating the kind of objective knowledge that is put at the center of WSYBO. Therefore, any notion of epistemic community that is articulated around group-based categories must explain how it would redirect the dynamics of exclusion and isolation typical of groups in order to ensure that the systemic epistemic benefits end up being positive. Are there any features of the epistemic communities exemplified in WSYBO that can minimize the corrosive nature of group dynamics? How to prevent the exclusion of perspectives? How to limit the polarization of opinions? How to minimize tendencies such as prejudices and biases that seem to prevent the articulation of objective knowledge? Does Frost-Arnold have in mind some cultural factors, organizational elements, rules or social norms that would minimize the effect of group dynamics on identity-related epistemic communities?

As I said before, the above offers an initial and sketchy route to question the epistemic potential that Frost-Arnold assigns to isolated epistemic communities organized around a shared identity. In what follows, I will try a second critical strategy, which involves exploring other possible varieties of identity-based epistemic communities. I think that keeping these possibilities in mind – and some of their negative outcomes and traits – can help us relativize Frost-Arnold's optimism in relation to epistemic potential of identity-based epistemic communities.

Let's go back to Frost-Arnold's example of a closed epistemic community organized around an identity marker:

> **(EC $_{good}$)** A blog for black women organized to share their experiences of discrimination in a specific work context. The group does not allow non-black women to participate (although it does allow 'lurking').

Closure to other perspectives helps members of this community to better articulate a particular perspective of oppression. Let's compare this community with two other varieties of epistemic communities organized along identity lines. Here is an example of what we could call an 'ugly' epistemic community:

> **(EC ugly)** A blog in which a group of mothers interact regularly about the risks of vaccines. The group allows any user to participate, although most interactions are structured around the values and beliefs shared by a stable group of mothers exchanging views on the topic in question.

And here is an example of a harmful or 'bad' epistemic community:

> **(EC bad)** A community of X users organized with the aim of marking as discriminatory or harmful any protest or critical content related to the interests of an oppressed group (p. 34).

Why does a group of black women who share their experiences of discrimination constitute a 'good' epistemic community, while a group of mothers who exchange views on the risks of vaccines offer an example of an 'ugly' epistemic community? And what separates these communities from those that are clearly harmful or 'bad'? Without aiming to be systematic (and with the intention of knowing Frost-Arnold's take on these cases) I think that several interesting things can be said about these cases. Some of these things could help us to understand the possible (and frequent) deviations that beset the 'good' epistemic communities highlighted by Frost-Arnold in WSYBO.

Let's start with the easy case. There is a clear sense according to which **(EC bad)** is a 'bad' epistemic community. It actively discriminates against members of other groups, favoring negative attitudes and emotions towards those people. These negative attitudes undoubtedly hinder access to the epistemic goods derived from contrasting opinions through open debate and inclusion. It could be claimed that in these cases the shared identity markers are framed 'negatively', that is, in opposition to other groups towards which discriminatory attitudes are actively promoted. These communities would be close to what Nguyen has labelled as 'echo chambers'. An echo chamber is a communicative structure in which other perspectives are let out *and actively discredited* (Nguyen 2018). When we keep this case in mind, it sounds reasonable to claim that, as least as a general rule, good epistemic communities must avoid fostering negative attitudes toward other groups. These negative attitudes could prevent the group from getting the epistemic benefits derived from interaction and openness. Good epistemic communities organized around a shared social identity would embrace identity positively, that is, as a criterion of belonging but never in opposition to another group.

Let's consider now the 'ugly' case. This one is a little more interesting. In (EC ugly) the group of mothers structures their interactions around a topic, the risks of vaccines, which constitutes the focus of a more general debate, present in other forums and contexts and largely governed by experts. An 'ugly' epistemic community does not actively exclude other groups, and in this particular sense it is clearly different from the 'bad' communities sketched above. However, this type of communities differs from (EC good) in a different sense, one that

seems very relevant and that would point towards another possible deviation to be avoided by a good identity-based epistemic community.

In Frost-Arnold's favored example, a group of black women organize their interactions with the goal of voicing or expressing their experiences in a certain work context. What characterizes these closed informational structures has to do in an important sense with the *security* they offer to their members to share these experiences - experiences endowed with epistemic value because they are peculiar and not mediated by relations of oppression, threats, etc. (Furman 2022). Keeping that general point in mind, the second deviation that I want to point out occurs between epistemic communities in which the identity marker is made explicit and puts us on the track of a forum in which we are going to learn things about the experience or experiences. of a certain group (EC $_{good}$) and those other communities focused on a particular topic or controversy (vaccines, climate change) and organized through non explicit dynamics of belonging and group affinity (EC $_{ugly}$). The risk with this second type of epistemic community is that now the usual groupish dynamics that we can find in any group can affect to how the group, as a deliberative system, would approach to the evidence and arguments surrounding the topic in question.

Dynamics of polarization and extremism, confirmation biases from dogmatically defended positions, policing of deviant opinions, and other similar phenomena could potentially affect the quality of the deliberation of these groups. Now it is no longer about expressing a perspective of oppression or a peculiar experience, the epistemic center of 'good' epistemic communities. Now it is about intervening in more general debates from a set of groupish dynamics inserted in closed informational structures.

Dan Williams has recently described these epistemic communities as groups in which a specific topic becomes a flag of group membership (see also Van Leeuwen 2023). Williams suggests that such groups, organized around certain 'identity-defining beliefs', exemplify processes of group epistemic deliberation of lower epistemic quality and negative aggregated effects. Williams mentions at least three of these processes: selection of evidence in line with the beliefs that define the group, dogmatic defense of those beliefs and what he refers to as 'markets of rationalizations' (Williams ms. See also Mercier & Sperber 2017. Gaitán 2024).

Above I asked if we can say something positive about the features that would separate 'good' epistemic communities organized around a shared identity from harmful or 'bad' epistemic communities (also sharing a common identity) and from those that I've referred as 'not so good' or 'ugly' epistemic communities. Now we can at least specify two general traits to be avoided:

— A good epistemic community organized around a shared identity must avoid defining itself negatively, that is, in opposition to other identities which are denigrated, insulted or stereotyped.
— A good epistemic community organized around a shared identity must protect a space in which experiences are expressed safely, avoiding taking a topic or debate as a flag of membership.

Bad epistemic communities organized around a shared identity are not necessarily bad because they limit interaction with alternative views. One virtue of WSYBO is precisely to remove this assumption, making space for a more positive view on the epistemic impact of a certain variety of closed epistemic communities. But epistemic communities organized

around a shared identity can be deviant for additional reasons beyond closeness. Here I've tried to highlight two of these reasons. They can be exclusionary and they can promote deliberative spaces where some topics or debates are behaving as flags of membership. There may be more negative features of identity-based epistemic communities, but I believe that the two I've just highlighted are especially salient in current digital contexts.

A precautionary note is in place. Maybe the image I am sketching here is too narrow to catch the political potential of closed (and good) epistemic communities. After all, there is a sense in which I am claiming that if they want to be epistemically fruitful, these communities must limit themselves to the articulations of experiences of oppression, avoiding wider political issues or debates, usually polarized and extremely divisive. This can be interpreted as conservative, even as regressive. But there are some tools in WSYBO that could help us to minimize this problem. We could secure the political potential of good epistemic communities by promoting lurking, or by establishing some deliberative satellites that could serve as a bridge between closed and good epistemic communities and the wider deliberative and political space (Squires 2002).

Conclusion

One of the virtues of WSYBO is the careful epistemic defense of closed communities organized around identity markers that it offers. In this commentary I have tried to specify the epistemic value of these communities, which according to Frost-Arnold has to do with the articulation of secure spaces in which oppressed groups are able to express their experiences. These shared experiences can influence the epistemic quality of the wider debate through various indirect means, from direct lurking to the articulation of regulated interactions with other communities or forums. However, the groupish nature of these communities also implies that their 'deliberative' focus must be kept very precise: it is advisable to avoid their interactions being articulated around specific themes or debates and in no case should they be defined negatively, as in explicit opposition to other groups. It is in the narrow margin where experiences of oppression are expressed where the value of these closed epistemic communities must be located.

References

Boyer, P. (2020). *Minds Make Societies. How Cognitions Explain the World Humans Create*, Yale University Press.

Furman, K. (2022). 'Epistemic Bunkers', *Social Epistemology*, DOI: 10.1080/02691728.2022.2122756

Gaitán, A. (2024). 'La psicología de las emociones políticas', en Gómez-Ramos, A. Velasco, G. (eds.). *Atlas de emociones políticas*, Madrid, Trotta

Habgood-Coote, J. (2024). 'Toward a Critical Social Epistemology of Social Media', in Lackey, J. McGlynn, A. (eds.). *Oxford Handbook of Social Epistemology*, Oxford University Press.

Mercier, H. Sperber, D. (2017). *The Enigma of Reason*, London, Penguin

Nguyen, C. T. (2018). 'Echo Chambers and Epistemic Bubbles', *Episteme*, doi:10.1017/epi.2018.32

Rini, R. (2017). 'Fake News and Partisan Epistemology', *Kennedy Institute of Ethics Journal*, 27 (52), pp. 43-64

Squires, C.R. (2002), 'Rethinking the Black Public Sphere: An Alternative Vocabulary for Multiple Public Spheres'. Communication Theory, 12: 446-468. https://doi.org/10.1111/j.1468-2885.2002.tb00278.x

Sunstein, C. (2009). *Republic 2.0*, Princeton University Press.

Van Leeuwen, N. (2023). *Religion as Make Believe. A Theory of belief, imagination and group identity*, Harvard University Press.

Williams, D. (ms). 'Identity-defining beliefs'