



# Through the Newsfeed Glass: Rethinking Filter Bubbles and Echo Chambers

Giacomo Figà Talamanca<sup>1</sup> · Selene Arfini<sup>2,3</sup>

Received: 24 May 2021 / Accepted: 5 December 2021 / Published online: 15 March 2022  
© The Author(s) 2022

## Abstract

In this paper, we will re-elaborate the notions of filter bubble and of echo chamber by considering human cognitive systems' limitations in everyday interactions and how they experience digital technologies. Researchers who applied the concept of filter bubble and echo chambers in empirical investigations see them as forms of algorithmically-caused systems that seclude the users of digital technologies from viewpoints and opinions that oppose theirs. However, a significant majority of empirical research has shown that users do find and interact with opposing views. Furthermore, we argue that the notion of filter bubble overestimates the social impact of digital technologies in explaining social and political developments without considering the not-only-technological circumstances of online behavior and interaction. This provides us with motivation to reconsider this notion's validity and re-elaborate it in light of existing epistemological theories that deal with the discomfort people experience when dealing with what they do not know. Therefore, we will survey a series of philosophical reflections regarding the epistemic limitations of human cognitive systems. In particular, we will discuss how knowledge and mere belief are phenomenologically indistinguishable and how people's experience of having their beliefs challenged is cause of epistemic discomfort. We will then go on to argue, in contrast with Pariser's assumptions, that digital media users might tend to conform to their held viewpoints because of the "immediate" way they experience opposing viewpoints. Since online people experience others and their viewpoints as material features of digital environments, we maintain that this modality of confronting oneself with contrasting opinions prompts users to reinforce their pre-existing beliefs and attitudes.

**Keyword** Filter Bubbles · Echo Chambers · Epistemic Bubble · Ignorance Bubble · Epistemic Discomfort · Digital Environments

---

This article is part of the Topical Collection on *Information in Interactions between Humans and Machines*

---

✉ Giacomo Figà Talamanca  
giacomo.f.talamanca@gmail.com

Extended author information available on the last page of the article

## 1 Introduction

The notion of filter bubble was prominently introduced in digital studies by entrepreneur and Internet activist Eli Pariser (2011). With the term “filter bubble,” Pariser designates a state of intellectual isolation determined by the preference algorithms that underlie contemporary web-based platforms such as Facebook and Google. Online users would be isolated, or “embubbled,” in the sense that they would consume content and interact with communities only *when in accordance* with their previous beliefs. In this way, they would tendentially be excluded from information sources and people in disagreement with their own perspective. They would, in other words, be limited to interact and consume information in an “ideologically safe” and unchallenged environment due to preference algorithms. In a similar vein, Cass Sunstein (2001, 2017; Sunstein and Vermeule, 2009) argues that the advent of web-based platforms favored the creation of intellectually secluded communities deprived of contrary perspectives, or *echo chambers*. For both authors, the diffusion of online platforms and of the algorithms that underlie their functioning lead to a significant reduction in political exchange between citizens of diverging opinions and increasing polarization within them, undermining the foundations of a fair and democratic society. Although the cause–effect relationship between the existence of both filter bubbles and echo chambers and the ruining of democratic discussions online might appear not just convincing, but potentially catastrophic, we aim to question this cause–effect relation that Pariser and Sunstein argue for.

In this paper, indeed, we will discuss and challenge Pariser’s notion of filter bubble and some of Sunstein’s claims on echo chambers, by proposing the re-elaboration of these notions in light of already established theories dealing with some epistemological downsides of fixating beliefs<sup>1</sup>. We argue that a distinction between technological and epistemological embubblements (the latter will be presented in Section 3) is necessary in order to account for a more complex relation between people’s epistemic status and their interaction with digital technologies. We will provide a novel reconceptualization of the concept of filter bubble that can account for such complex relation and that is more fitting with existing empirical research on online information consumption behavior and the intellectual isolation Pariser and Sunstein are concerned with. Our reconceptualization will propose that such intellectual isolation does not derive from the activities of the algorithm alone, but rather from the interaction between the user’s beliefs and cognitive profile and the platform’s interface, which lacks the contextual norms and socio-emotional cues that would make facing contrary viewpoints more functional.

---

<sup>1</sup> A terminological clarification here might be necessary. With “the fixation of beliefs” we do not intend to refer to obsessive or irrational attachment to particular ideas. Instead, we, as the authors we will refer to, follow the pragmatic tradition, opened by Charles Sanders Peirce, which presents the state of belief as, simply, a stronger description of belief with respect to the one adopted by commonsense: “[the belief] is a calm and satisfactory state which we do not wish to avoid, or to change to a belief in anything else.” (Peirce, 2011, p. 41). In this account, people fixate on some beliefs because “we cling tenaciously, not merely to believing, but to believing just what we do believe.” (p. 41)

The paper will proceed as follows. In Section 2, we will present the notion of filter bubble in Pariser's formulation. Specifically, we will discuss its relation to the notion of echo chamber (Sunstein, 2001; Sunstein and Vermeule 2009) and we will argue that both insufficiently provide explanations for the phenomena they aim to account for since there is a lack of empirical evidence that supports them and, they generally favor misunderstandings regarding the impact of technological developments on society. In Section 3, we will lay the grounds for a reformulation of those concepts, by reviewing previous theories of epistemological "embubblemment," elaborated by Woods (2005), Magnani (2011) and Arfini (2019).<sup>2</sup> For these authors, people are normally subjected to different kinds of embubblemment, intrinsic to some of their cognitive processes, such as the "fixation" on certain beliefs, their tendency to follow moral norms, and their underestimating their own ignorance. In Section 4, we will revise the notions of filter bubble and echo chambers as "techno-epistemological" ones, in the light of the previously laid out theories. The definition of such mechanisms will not be a consequence of a reduction of a technological embubblemment to an epistemological one (similar to the one already performed by Pariser): we will indeed argue that the filter bubble (as well as online echo chambers) should be understood as the result of people's "natural" embubblemment—understood as a series of default and ultimately inescapable conditions of human agents that we will present in Section 3—while interacting with digital environments. Specifically, Internet users often stumble upon opinions and viewpoints that contradict theirs, and it might strengthen their own beliefs due to what we will define as the "immediacy" of these encounters. This immediacy is the result of a characteristic of many mainstream digital platforms, known as context collapse—which we will also comment on subsection 4.2 (Marwick and boyd, 2011; Vitak, 2012; Costa, 2018).<sup>3</sup> So, we argue that the unmediated way users see opinions of other people that openly contradict their own views prompt them to a form of cognitive rigidity—i.e., the tendency to fixate on one's own beliefs and to take them as true until proven

<sup>2</sup> We acknowledge that, in a recent publication, Nguyen (2020) utilizes the notion of epistemic bubble to reflect on the formation of echo chambers online and on information consumption behavior. We want to remark here that the two notions are very different: while Woods' notion intends to capture a phenomenological and structural character of having beliefs (which is not a psychological phenomenon but may have psychological consequences), Nguyen uses this notion to refer to "a social-epistemic structure which has inadequate coverage through a process of exclusion by omission" (143). Nguyen's notion is closer to our notion of "merely informational environments" (see sec. 4.1, p. 19) as spaces that enable highly personalized ways of managing information and possibly (not necessarily) excluding or omitting other perspectives.

<sup>3</sup> Importantly, we acknowledge that the overall mechanisms of digital platforms as business models are mediated by the service providers' intention to make a profit; however, this is not the relevant sense in which we talk about (un)mediation. Firstly, while we realise the importance of the capitalist logic of the service providers in the design of digital platforms, for the purpose of our paper we are focusing on the relation between individuals and artefacts, and how this relation transforms the individual's cognitive processing (see e.g. Clark and Chalmers 1998; Clark 2003, 2008; Verbeek 2008; Clowes 2015; Smart 2017; Ihde and Malafouris 2019; Baber, Chemero and Hall 2019; Rosenberg 2019). Secondly, while we acknowledge that the transformative impact of these technologies on cognitive processing can be understood as a form of mediation, we are not following Ihde's (1990) or Verbeek's (2005, 2011) theory of technological mediation specifically. With the word "immediacy" and "unmediated" we are specifically talking about a manner of experiencing information in online environments that is determined by some characteristics of digital platforms - most prominently, context collapse. We examine the meaning of this immediacy in detail in section 4.2.

contrary<sup>4</sup>—instead of openness and debate. These considerations will shift the focus of “the filter bubble” from technology alone to the interaction between believers and technology, leading to the conclusion that, in a sense, the filter bubble is “already in your head.”

## 2 The Filter Bubble: Relevance and Shortcomings

### 2.1 The Filter Bubble Thesis

Pariser uses the term filter bubble to describe the impact of personalized search algorithms on Internet platforms such as Google and Facebook. The emergence of these developments, which occurred in the late 2000, marked the advent of what is called “the Web 2.0” (O’Reilly 2007), which enabled Internet-based platforms and services to be more focused on interactivity, social networking and user-generated content. Indeed, Pariser argues that the filter bubble is a fundamental part of Web 2.0 since it results from the personalization process enacted by web-based platforms in order to (allegedly) enhance user engagement, interconnectivity, and content consumption. He argues that “[t]he new generation of Internet filters looks at the things you seem to like ... and tries to extrapolate. They [i.e., the algorithms] are prediction engines, constantly creating and refining a theory of who you are and what you’ll do and want next. Together, these engines create a unique universe of information for each of us—what I’ve come to call a filter bubble—which fundamentally alters the way we encounter information online” (Pariser 2011, 9).<sup>5</sup>

A downside of this technological development is, according to Pariser, the intellectual isolation of Internet users: “you’re the only person in your bubble. In an age when shared information is the bedrock of shared experience, the filter bubble is a centrifugal force, pulling us apart” (p. 10).

This condition of intellectual isolation is determined by the following process:

- 1) Users engage with certain content on the Web, through search engine services like Google or SNSs such as Twitter or Facebook;
- 2) The algorithms underlying those platforms identify the content preferred by users;
- 3) Algorithms will provide new contents based on users’ preferences, that is, on their previous engagements with specific matters instead of those they did not engage with.<sup>6</sup>

This mechanism is in line with common understandings of the functioning of intelligent software: the algorithm chooses a set of possible actions to be presented

<sup>4</sup> See footnote 1 at page 2 of this article.

<sup>5</sup> It is relevant to note, as Bruns (2019) points out, that while there is a generalised understanding in the literature of what a filter bubble is, Pariser (2011) does not provide a univocal definition, and definitions vary slightly across publications of different authors. The same occurs with Sunstein’s notion of “echo chamber” that we introduce in section 2.2.

<sup>6</sup> At best, preference algorithms will provide content similar to that previously consumed.

for users, who then choose one of the possible courses of action and the algorithm can further favor engagement by providing other possible actions based on the user's past choices (Burr et al., 2018). The logic of the filter bubble thesis is furthermore grounded in the service provider's economic gain. As Srnicek (2017) points out, the aim of platforms such as Facebook, Twitter, or Google—i.e., digital infrastructures that enable agents to interact and act as intermediaries between customers, advertisers, service providers, producers, suppliers and physical objects and events (43)—is the gathering of data regarding their users' preferences and activities, which is in turn the online service provider's source of revenue. This is especially so for advertising platforms who sell the extracted data through users' online activities to advertisers—and it is this kind of platforms that the filter bubble thesis primarily targets, services like Google, Facebook, or Twitter. If these companies aim at data gathering, and if the more data they get the more they can capitalize, the filter bubble thesis seems to fit very well in this dynamic—for, at least at a first glance, the quickest way to get user engagement would be to provide content they (seem to) like or have engaged with before.

Pariser provides an example in order to show the dangers this mechanism can entail. Suppose you have friends on Facebook of different political orientations; and suppose that, as you are more inclined to consume content on the left side of the political spectrum, you visualize and engage with posts by those friends whose political views are closer to yours. What happens, Pariser argues, is that the algorithms underlying Facebook's newsfeed will take note on the fact that you engage more with content on one side of the political spectrum and less with the other: from there, the algorithm will provide you with content similar to what you previously consumed, by presenting on your newsfeed more posts on the left side of the political spectrum and not from the right.<sup>7</sup>

To summarize, preference algorithms, Pariser argues, accommodate users by providing content consistent with what they consumed in the past: in the example he provides, the user who mainly engages with news consistent with her political views will not receive those from contrary or dissimilar political standpoints. The preference mechanism enacted by the algorithms of contemporary web-based services would lead to an informational embublement: algorithms would filter information inconsistent with users' online engagements, so that they would not encounter information opposing their previously shown convictions and beliefs.

It is easy to see the detrimental side of this scenario, *were it true*. If indeed the algorithmic structure of web-based platforms leads to an informational embublement (personalized for each user's preferences and engagements) users would not encounter standpoints and information that challenge their standing attitudes and beliefs. If this were to be the case, users of digital platforms would likely become more hardened in their beliefs and political attitudes; this radicalization would, in turn, lead to increasing polarization between different standpoints. The lack of interaction with people of different

---

<sup>7</sup> It is worth noting that one of the aims of preference algorithms is what is called online profiling, the recording of online behavior of internet users in order to promote product advertisement. There is a somewhat consistent literature on how profiling constitutes a harm to personal autonomy (see Hildebrandt 2008; Büchi et al., 2019).

standpoints would lead to a complete lack of common ground where a minimal degree of mutual understanding would be possible, a necessary feat for a functioning democratic society. Specifically, if a democracy properly functions through the existence of diverse standpoints and the dialogue between them, the creation of ideologically segregated communities would hinder the very capacity for democratic debate. The relevance of this topic is apparent since, beyond Pariser's account, various authors have tackled and discussed the relationship between platform capitalism's algorithms and online polarization (Gurumurthy & Bharthur, 2018; Marciano, et. al. 2020; Riemer & Peter, 2021).

As Pariser argues, this mechanism would have deep implications that do not just involve political interaction, but also the state of mind of individual citizens and the general social and political role of these platforms in the years to come. Or as he puts it: “[w]hile the Internet offers access to a dazzling array of sources and options, in the filter bubble we’ll miss many of them. While the Internet can give us new opportunities to grow and experiment with our identities, the economics of personalization push toward a static conception of personhood. While the Internet has the potential to decentralize knowledge and control, in practice it’s concentrating control over what we see and what opportunities we’re offered in the hands of fewer people than ever before” (218). The conclusion, it is argued, is that web-based platforms have severe consequences for democracy and individual and collective well-being, since the preference algorithms that underlie them have the undesirable effect of secluding and polarizing users depending on their attitudes and beliefs.

## 2.2 Filter Bubbles and Echo Chambers

The notion of filter bubble and Pariser's thesis regarding its existence and impact is not a solitary instance when it comes to theorizing the (potentially catastrophic) impact of digital technologies in society. The notion of filter bubble is in fact closely associated with what Cass Sunstein (2017) calls “echo chambers.” Sunstein argued that users of the “Web 2.0” would have the “power of personalization, or to create gated communities” (Sunstein, 2017 p. 5) composed by people sharing identical or similar beliefs and convictions. Only previously shared and accepted opinions would be shared in these communities, and beliefs and attitudes challenging those opinions would be excluded. The term “echo chamber” denotes a social network (i.e., a community of people with social ties with one another) who share a (set of) opinion(s) while not interacting with opinions and viewpoints that would contradict them.<sup>8</sup> Whereas echo chambers can actually exist in

<sup>8</sup> It is appropriate, here, to make some terminological distinctions. With the term *social network*, we mean to describe a group of social actors interacting with each other, either offline or online. With the term *social networking site*, we mean to describe online platforms (such as Facebook or Twitter) where users interact with each other - in other words, digital platforms where social networks can take form. With the term *online communities*, we mean social networks that specifically form in digital platforms - groups of actors interacting with each other online. To further clarify: for Pariser, filter bubbles are not social networks, but the way the platform's algorithms affect an individual user's chances to encounter information. In contrast, echo chambers are social networks which are affected by either the choice of each individual within the network or by algorithmically generated informational seclusion (see section 2.3). Along with the concept of filter bubble, in the remainder of the paper we criticise the existence of echo chambers that are (allegedly) created through algorithmic choice and not users' choice.

offline settings,<sup>9</sup> on the Internet they would skyrocket in membership numbers and have a much more significant impact than they would in their offline counterparts.

Despite some similarities between filter bubbles and echo chambers, there are two clear differences between the two notions. Firstly, filter bubbles describe forms of intellectual isolation exclusively caused by algorithms (which inferences are based on users' choices). In contrast, echo chambers are enacted by users themselves. They can exist in offline settings, but (Sunstein argues) are widened in online ones: users of digital technologies would interact only with people who share their same beliefs and values and exclude diverse perspectives (also thanks to algorithmic intercession).

Secondly, the two notions apply at different levels of abstraction: Pariser's filter bubble aims at describing online information consumption and exposure from an individual user's perspective, while Sunstein's echo chamber applies to interaction within a (online) community. As Bruns (2019) puts it, the notion of echo chamber applies to a social network in its entirety, as it comes into being when some users choose to connect with others, shaping up a group and excluding outsiders. The more the network's borders are sharply distinct, and the more connections are created within it, the more isolated the network is from outsiders and their potentially contrary standpoints. In contrast, filter bubbles are theorized to occur at an individual level. Specifically, the more consistently users consume specific information instead of other and the more they communicate with certain users (who share the same interests and opinions). In turn, it would be more likely that the users' own views and information will circulate within their network and be confirmed, rather than any view or information from the outside.

To summarize, both these notions share an important assumption, which will be closely examined throughout the course of this paper. Specifically, both these theses share a specific view on the quality of digital technologies' impact on society. Both echo chambers and filter bubble are seen as a novel dynamic that had a significant impact in society as technologically determined (specifically the filter bubble) or incentivized (specifically echo chambers). So, their most relevant common trait, *would they actually exist as Pariser and Sunstein intended*, would be their being generated and potentiated by digital technologies, independently of the wider societal context where they are introduced.<sup>10</sup> It is this technological character that turned a great deal of attention from both academia and the general public to the negative ways digital technologies affect daily life.

---

<sup>9</sup> For instance, a group of people who would consume television based news from one channel only; or a physically isolated community of people who share the same set of views on a certain topic (e.g. all of whom are Catholic) are examples of offline echo chambers.

<sup>10</sup> To be clear, as we will argue in the next section, we think this sentiment, shared by Pariser and Sunstein, is fundamentally misguided, for it assumes a form of technological determinism that we do *not* endorse.



### 2.3 Four Problems with the Filter Bubble Thesis: Toward a More Inclusive Understanding

After having introduced the notion of filter bubble and its relation to the notion of echo chamber, we should briefly survey three criticisms that both these ideas faced in the years succeeding their introduction. We will also advance a fourth critical comment regarding these theses, which will give shape to our argument in the rest of the paper, helping us re-thinking the filter bubble and the echo chambers hypothesis as a by-product of a more articulated epistemological situation.

The first amply accepted criticism regards a fundamental aspect of both these notions: the idea of *informational seclusion*. Both Pariser and Sunstein argue that Internet users will tend to find and engage with information and other users that agree with their views. For Pariser especially, such seclusion occurs thanks to the platforms' algorithms and the way they provide information to individual users. In contrast, for Sunstein such seclusion can be generated both by user choice and by algorithmic recommendation—possibly at the same time—and affects an entire social network of Internet users. However, with the exception of research from Flaxman et al (2016), Quattrociochi et al. (2016), and Wollabæk et al. (2019), the wide majority of empirical research found very little evidence of algorithmically generated informational seclusion. In other words, people online not only do see, but also engage with pieces of information that oppose their previously and currently held beliefs.<sup>11</sup>

In the case of Google, previous searches and geographical location were not found to affect the search engine's results in any way when it comes to news and potentially polarizing issues<sup>12</sup> (Haim et al., 2018; Krafft et al., 2019; Nechush-tai and Lewis, 2019; Cardenal et al., 2019). In the case of Facebook, most users seem to find and engage with people of different opinions (Beam and Kosicki, 2014; Beam et al., 2018) and to be exposed to a variety of different standpoints (Bakshy, et al., 2015; Fletcher & Nielsen, 2018a, 2018b; Fletcher et al., 2020). In fact, members of an online community who share certain views might actually seek out other users with divergent beliefs (Smith and Graham, 2019). In the case of Twitter, some politically savvy users were found to be “embubbled” within the

<sup>11</sup> It is worth pointing out that the wide majority of these studies, with the partial exception of Fletcher et al. (2020) only considered WEIRD (Western, Educated, Industrialized, Rich and Democratic) countries, and are not necessarily generalizable to the entirety of Internet users around the globe. However, neither Pariser nor Sunstein seem to consider non-WEIRD countries in their theories. More in general, this collection of studies we mention does not aim to prove these theses wrong beyond reasonable doubt, but more simply to question certain aspects that characterize them and might be changed.

<sup>12</sup> This is not to say that localisation does not affect Google searches at all, a fact there is clear evidence for (Ochigame and Ye 2021). When a user activates geolocalisation on their devices and access to those data is granted to the search engines installed in those devices, at least some of the sources will be affected by geolocalisation. For instance, Google results for the term “God” will vary depending on the location of the user. Similarly, searching for a term such as “sushi bar” with active geolocalisation will tendentially lead to a list of sushi bars in the user's vicinity. What we are concerned with here (and what Pariser and Sunstein are concerned with) is news-generated information or potentially polarising issues; and we did not find research corroborating the idea that geolocalisation, demographics or previous searches affects the visualization of that kind of information on Google.



ideology of the online community they belonged to, in contrast to more casual users in the same network (Garimella et al., 2018; Williams et al., 2015). However, these online communities are not close-knit: networks of users might cluster around shared interests and topics, but their members are still in contact with other networks (Bruns et al., 2017) and sometimes they actively seek groups with opposing standpoints to engage them in debate (Yardi and boyd, 2010). Finally, some of the studies claiming an occurrence of informational seclusion typical of filter bubbles and echo chambers had somewhat contradictory findings: users were found to be polarized and more likely to conform to preexisting beliefs *and at the same time* to be exposed to opposite standpoints (Flaxman et al., 2016; Quattrociocchi et al., 2016; Wollabæk et al., 2019).<sup>13</sup> All these studies seem to show that Internet users find and engage with standpoints different from their own, rather than only with content conforming to their views.

The second criticism (which we direct to Pariser's argument specifically) regards the presuppositions of what kind of information people engage with online, and why. Pariser seems to imply that the wide majority of Internet users will engage in politically oriented searching behavior, making a somewhat strong presupposition on *user's preexisting attitudes as politically oriented*. The argument presumes that users are mainly motivated by political engagement while using web-based platforms. That is not necessarily the case (Fletcher et al., 2020; Bruns, 2019): actually, people might, more often than not, be incidentally—not intentionally—exposed to politicized information online, while they might mainly use the Internet for other purposes such as entertainment or sociality.

The third criticism is constituted by the general attitude manifested by Pariser and Sunstein toward the social impact of digital technologies, an attitude describable as *technological determinism*. This is the idea that “the technology prevalent in a society will drive the behaviors of its citizens and hence the social structure as well as cultural values” (Dutton et al., 2017) As Bruns (2019) and Fletcher et al. (2020) point out, the inherent problem of the notion of filter bubble—and of echo chamber, to a lesser extent—is the appeal to “purely” technological dynamics to explain and warn against negative or unexpected socio-political developments. They both point out that these terms entered everyday discourse after critical events such as the election of US president Donald Trump and the results of Brexit referendum in England. Bruns argues (2019, 115–121) that the rise to prominence of filter bubbles and echo chambers as explanatory concepts for these political developments might be more of an explanatory shortcut than an accurate account. They would help shift the blame for societal problems to technology, rather than actually understand those problems in their wider context. Moreover, whereas Big Data companies such as Facebook could definitely improve their standards when it comes to the ethical and societal implications of their corporate decisions (Grimmelann, 2014; Cadwalladr, 2018; Lapowsky, 2019), to argue that any political development rises from black-boxed algorithmic dynamics overshadows already present problems at a

---

<sup>13</sup> These studies will be reconsidered more closely and re-evaluated in section 4, when we will reformulate the notion of filter bubble.

societal level. The technological determinism assumed by Pariser and Sunstein, in other words, might oversimplify the impact of digital technologies in our society. And if that is the case, then, research trying to understand these dynamics should expect that the relevant processes characterizing the impact of these technologies in everyday life should account more than the effect of algorithms over how people consume information.

Thus, a fourth criticism should be advanced at this point, discussing in detail whether and in what sense the filter bubble and online echo chambers should be considered socio-technical problems, after pondering on the justified accusation of technological determinism that depicts Pariser's and Sunstein's approaches. Taken as epistemological arguments (and not as sociological ones), their approach would aim at explaining belief fixation, polarization, and radicalization in social networks with the creation and diffusion of particular technological systems. Without arguing against a correlation between the two events, we argue against the possibility of speaking of causation between the two.

To sustain our position, we can comment on one experiment that tries to consider the relation between technology and their sociocultural context. The experiment was conducted by Davies (2018), who confronted two groups of British high school students from different social classes, in order to investigate whether Google would provide different results depending not just on their previous searches, but also on their level of media literacy and awareness about the research engines. Whereas the participants were relatively few, he did confirm his original hypothesis: the informational embublement theorized by Pariser occurred only in some cases. In these cases, the embublement was heavily determined by the persistence of the subjects' previous beliefs in front of results that seemingly disconfirmed them, which also depended on the subjects' varying degree of expertise about how to look for information online. In one case, a student was asked "Should we [England] have more or fewer restrictions on immigration?"; and having already the opinion that more restrictions should be made, he typed search terms such as "how much benefits do immigrants get" and "immigrant mansions," instead of "reasons for immigration" or "immigration economic implications" (p. 649–650). Using different search terms that are supposedly related to the same topic can lead to very different information and perspectives (Borra & Weber, 2012). For this reason, Davies proposed, and we agree, that we should rethink filter bubbles and—we add—online echo chambers as socio-technical problems, whereby not only digital technologies but also pre-existing attitudes constitute an informational embublement. Thus, if technological determinism falls into a sort of "correlation is causation" fallacy (which mistakes the radicalization and polarization in social media as an effect of the personalization algorithms instead of contemplating much more complex and old-school reasons), then we are left with a question that demands a more encompassing answer. In which kind of relation should we think of the sociocultural intellectual seclusion and the technological one?

Following the idea that an epistemic (and not simply informational) seclusion might be already in the user's head *before* her engagement with digital technologies, in the next section we will discuss some theories that have been advanced regarding which kind of embublement can affect people's attitudes toward the formation of

new beliefs; and why at most digital technologies might play a role in either mitigate or radicalize them—in a reflecting, instead of determinist or symbiotic, kind of fashion.<sup>14</sup>

### 3 Epistemic (and Other) Bubbles and Epistemic Discomfort

The possibility that the agent could hold on and share beliefs and positions with unjustified confidence, even arrogance (Tanesini, 2016), has been one of the key topics of contemporary theoretical and social epistemology before the emergence and diffusion of digital technologies (Cassam, 2018). Now the discussion has only an ampler range of applications. In particular, the so-called “bubble theses” (Arfini, 2019) are arguments that aim at discussing how some common cognitive states (such as beliefs, doubts, and moral stances) can both favor the epistemic growth of the human agent and become sources of delusion for her. The theoretical assumptions of the bubble theses are few and simple: a) human beings are fallible agents that fixate beliefs to act into uncertain situations; b) from a phenomenological perspective, believing to know something (which can be seen as fixating a belief) and knowing the same thing appear the same; c) hence, confidence regarding one’s own beliefs is intrinsically highly fallible and may delude the agent on the robustness and legitimacy of her positions. Thus, we turn to review the “bubble theses” because we maintain that they provide an excellent foundation for re-elaborating the concept of filter bubble, by grounding it in the intrinsic cognitive and epistemic limits of human beings, and how these constraints are affected by the interaction with digital platforms, intended as artefacts.<sup>15</sup> The name and scope of the bubble theses that provide such grounding are the following:

- 1) The *epistemic bubble*, which has relevant implications when considering how people gain knowledge and feel the need to do so;
- 2) The *moral bubble*, which affects how people morally evaluate each other’s actions;
- 3) The *ignorance bubble*, which has an impact on how people estimate their ignorance.

---

<sup>14</sup> While our complete re-elaboration of the notion of filter bubble will be done in section 4.4, we want to already clarify here that our understanding of filter bubble as a sociotechnical seclusion is related but distinct from other cognitive mechanisms such as motivated reasoning or motivation biases. Rather, we intend to re-elaborate the notion of filter bubble as an epistemological and affective state occurring while interacting with digital technologies that can in turn trigger psychological responses, biased and motivated rationality processes.

<sup>15</sup> With the terms “limits” and “constraints” we are not targeting the faults of some agents. Rather, we are specifically talking about limits - as constraints - that derive from the inevitable situatedness of our potential knowledge and general existence. We simply cannot have a “God’s eye view” on whatever issue we are dealing with because our brains have limited space and our bodies are inevitably situated within specific sociomaterial settings. That being said, of course some people might be more reflexive, curious or open-minded than others, but those epistemic virtues are limited by our natural boundaries as situated organisms. The bubble theses aim to describe some of these unavoidable constraints, and not some defects agents are culpable for.

In this part of the article, after describing these situations in detail, we will present how taking them into consideration can be of use also when reflecting on problematic human-to-human interactions in digital environments.

### 3.1 The Epistemic Bubble: The Categorical Shortcomings of Belief

Woods's (2005) notion of epistemic bubble originates from a few somewhat truisitic considerations: human beings ordinarily make many errors regarding what they know and what they think they know. However, their tendency to formulate accurate beliefs is generally taken to be a relevant contributing factor for the sophistication and success of our species (Clarke, 1990; Griffith and Wilkins, 2010; McKay & Dennett, 2009, Woods, 2005). Indeed, when knowledge is necessary for action, and when the agents are aware that they lack the necessary knowledge to act, they become cognitively irritated<sup>16</sup> and develop the urge to relieve this irritation by elaborating a belief (Peirce, 2011). The possession of beliefs (in best case scenarios, of knowledge) is in principle pleasurable: agents, by formulating a belief that they think is true (or reliable enough to act upon it), relieve the psychological and practical tension derived from not knowing what to do in particular circumstances. However, this situation creates a phenomenological issue: from the first person perspective the agents do not think they *believe* something, but that they *know* the same thing. Hence, it is impossible for the agent to discern, phenomenologically and in the present moment of the belief formation, whether the agent knows or *simply believes* to know something. Such discernment is possible only in the aftermath of the reliance on the belief, either through the agent's experience or from other people's feedback.

Knowledge is, on the one hand, elaborated through beliefs since, of course, knowing something implies believing to know the same thing. However, because belief and actual knowledge are phenomenologically indistinguishable, the formulation of beliefs can actually prevent their verification as knowledge (because the agents already *believe* that they *know*). In this sense, "belief is both a condition of knowledge and an impediment to its attainment" (Woods 2005, 739).

Of course, this embubblement does not make it impossible to gain actual knowledge. Human beings' openness to third-parties' feedback, as well as their capacity to gain new information, can support the adjustment of beliefs to reflect more accurately the state of affairs. This, at the same time, does not make the revision of beliefs an escape route from the epistemic bubble. Rather, the continued revision of beliefs is implied as a shift to an epistemic bubble to another. As Woods describes it, even knowing about the epistemic bubble does not make us immune to it: "Although a cognitive agent may well be aware of the bubble thesis [of the structural incapacity to distinguish knowledge and belief] and may accept it as true,

---

<sup>16</sup> With the term "cognitive irritation" we mean a psychological (and specifically affective) state that is intrinsically and specifically related to the manner agents manage information and form beliefs. This term is already used by Hookway (1998), Pierce (2011) and by Woods (2005) with this meaning.

the phenomenological structure of cognitive states precludes such awareness as a concomitant feature of our general cognitive awareness” (743).

So, to be aware of one’s own epistemic bubble does not, in principle, support a way to “escape” it. Unless the agents *become aware that* (as they formulate *another belief* about) there is pressuring evidence or trusted social feedback in favor of the contrary, they will maintain that they know what they think they know. In other words, the feedback provided by not just other viewpoints, but also conventions, common sense and social norms is also a condition for furthering knowledge and understanding and at the same time an obstacle to its attainment.

### 3.2 The Moral bubble: the Systematic (Mis) Representation of Violence

The sociocultural aspects of the epistemic bubble have been especially investigated by Magnani (2011), who specifically reflects on the human limitations in the understanding of violence. Magnani argues that the problem of the epistemic bubble encompasses also moral beliefs, our opinion on what actions qualify as violent (and so as moral or immoral). The ordinary tendency to maintain beliefs and take them as legitimate, highlighted by the epistemic bubble thesis, is even more evident when it comes to moral beliefs (Holt et al., 2009). The experienced legitimacy for maintaining a moral belief is not just motivated by the perceived state of affairs, but also by variables dependent on the social and cultural context of the agents involved (Sommers, 2009). The experience of validity and legitimacy characteristic of moral beliefs is not due to the need to describe a state of affairs in the world, but to the expectation of how the world is supposed to be (Boyd & Richerson, 2001). Magnani points out that moral norms, options, and orientations so constrain our perspectives in exclusion of others: I might have a belief regarding certain practice being violent (and so immoral), but this belief and its sources (such as my education, social class, and moral norms) prevent me from the possibility of unbiasedly evaluating the presumed violence of the target practice.

An example can help clarify the scope and implications of the moral bubble. The veil (*hijab*) worn by Muslim women might be seen by Western atheists or people of Christian descent as symbolic of a form of oppression toward women (Ruby, 2006). However, part of the reason why Muslim women ought to wear a hijab is that this piece of garment is meant to help them avoid harassment and undesired attention from other men (Gabriel & Hannan, 2011). In a case such as this, the very same practice can be seen as a perpetration of violence on the one hand and as the exact opposite on the other: the moral beliefs of each perspective on this norm is diametrically opposed.

Moral beliefs, therefore, constitute what Magnani calls *moral bubbles*.<sup>17</sup> The notion of a moral bubble shows that our epistemic capacities not only limit our discerning between knowledge and mere belief, but also the discernment between

---

<sup>17</sup> The most straightforward (and sufficient for the purpose of this paper) way to understand this relation would be the following: every Moral Bubble is also an Epistemic Bubble, but the reverse does not stand.

legitimate and illegitimate actions. This discernment is, furthermore, reinforced by, if not outright grounded in, the moral community the subject belongs to: the possession of moral beliefs is greatly determined by the norms and values shared by the community. To follow a system of moral beliefs and values, which can vary across cultures, has a cooperative function and can be seen as evolutionarily advantageous (Sommers, 2009). However, it is at the same time a system that allows violent behaviors through practices of norm enforcement such as punishment, sanctioning or mobbing. It is due to this structural feature that human beings are incapable of recognizing—or they morally justify—the presence of violence in their own moral beliefs (and related moral behaviors).

### 3.3 The Ignorance Bubble: the Underestimation of the Unknown

Arfini (2019), elaborated what can be called a foil of Woods's epistemic bubble, that highlights even more the structural limitations of human cognition. Just as human beings are phenomenologically incapable of distinguishing between what they know and what they believe they know, they are also incapable of discerning what they do not know and what they think they do not know. In other words, human beings are naturally prone to underestimate their ignorance, and to think they know what are the things they do not know. If the incapacity to distinguish between knowledge and belief is at the core of the epistemic bubble, the incapacity to distinguish between our actual ignorance and what we think is the extent of our ignorance (our doubt, uncertainties, etc.) is the core of the *ignorance bubble*. And while the ignorance bubble is a complementary aspect of the epistemic bubble, its extent can be considered much wider.<sup>18</sup>

When we are stricken by doubt, by the awareness that we do not know something, we are well aware of it due to the irritation it brings us as cognitive agents, and requires effort and reasoning to solve it. However, doubt only involves the amount of ignorance we are aware of and it is almost physiologically impossible to be aware of the actual extent of our ignorance. Our irritation would make us motionless and incapable of proper everyday action. Author's reflections point out, in other words, that a cognitive limit has an impact on how exactly we can estimate our ignorance.

Of course, the human tendency to underestimate their ignorance is not a novel research issue in psychology and philosophy of mind. A notorious example is the Dunning–Kruger effect (Dunning and Kruger, 1999, 2002; Dunning 2011): that is, the tendency of people to overestimate their competence when they have little

---

<sup>18</sup> The concept of the Ignorance Bubble is notably in tune with various currents that now compose the so-called Ignorance Studies (Gross and McGoey 2015; Peels and Blaauw 2016; Proctor and Schiebinger 2008; Sullivan and Tuana 2007) in which sociological, psychological and philosophical reflections on the concept of ignorance now have found a venue to discuss its complex and interesting features.

in certain domains. However, while the Dunning–Kruger effect describes cognitive shortcomings of agents in particular circumstances (developing a skill set, for example), the ignorance bubble, just as the epistemic bubble, describes a structural feature of human cognition, that is, in a sense, practically motivated: a lot of the things we do not know are not useful to carrying out our everyday business.<sup>19</sup> To put it with a blunt example, the knowledge of the diameter of the Earth or of the subatomic composition of Mendeleevium is not intuitively handy for a philosophy student. Nor the human brain is computationally capable of processing the entirety of knowledge coming from all the existing scientific disciplines. Ignorance needs to be accepted as a pervasive feature of everyday life, as human beings, the subjects of ignorance and knowledge, are situated beings with situated interests and goals.

The ignorance bubble thesis, however, has some somewhat more disquieting implications than Woods' original argument. When people engage in their everyday business, while they might feel the pressure to revise what (they think) they know, the same often does not necessarily hold for what they (think they) do not know. The agent's ignorance is understandable as a massive frame that she cannot help ignoring. The ignorance bubble represents, therefore, a much more extensive form of embubblement than the epistemic bubble, where the success of the agent's everyday and situated practices impedes her from realizing the extent of her ignorance and from questioning how much she does (not) know.

### 3.4 The emotional Side of the Bubbles: Epistemic Feelings and Discomfort

A stable feature of all these bubbles is their practical function in the everyday experience of the agents, that allows them to believe they have knowledge at their hands, to feel that they are not behaving violently, and to avoid feeling crushed by the depth of their ignorance. At the same time, what allows us to discuss and describe the function of these bubbles is the fact that in some circumstances agents do need to confront the fact that their knowledge is not as sound as they thought, their behavior might be not morally accepted by others, and their ignorance is much deeper than they predicted. In these cases, what has an impact on the agents is not just the epistemic or moral acknowledgement of a situation, but also related emotional responses which are studied in metacognitive literature with the term *epistemic feelings*.

Epistemic feelings are elements of the emotional spectrum (Arango-Muñoz, 2014) with meta-cognitive functions and which affect cognitive abilities, states, and decision making (Sousa, 2009; Evans, 2008; Terpe, 2016). Examples of epistemic feelings that we experience in everyday situations are the feeling of knowing or the tip-on-on-the-tongue feeling. Doubt, in this sense, is irritant (also called epistemic anxiety in psychological literature—Hookway, 1998), because it is the acknowledgment of one's own ignorance. As we already discussed, when we experience doubt

---

<sup>19</sup> It is also worth pointing out that, because the ignorance bubble encompasses a structural aspect of human cognition and knowledge-formation, and while the Dunning-Krueger effect applies only to specific circumstances, evidence that seem to falsify the original findings by Dunning and Krueger (Ackerman et al. 2002; Nuhfer et al. 2017) does not undermine the ignorance bubble thesis.



we want this sensation to end as soon as possible: agents can end this state of irritation by acknowledging the doubt for what it is, and by making the cognitive effort to get to know the state of affairs that is the source of doubt. However, sometimes making this effort is not feasible, either due to a lack of previous information or due to a lack of motivation, and agents might simply downplay their ignorance in order to avoid the irritation brought about by the awareness of it. The Dunning–Kruger effect is an instance of the way human beings devalue, often unintentionally, the width of their own ignorance. And the downplaying of one’s own ignorance is not unmotivated, nor intrinsically vicious.

For the purpose of this paper, we will highlight the experience of not-(yet)-acknowledged doubt: the situation in which people believe to know something instead of actually knowing it, and they actively try to avoid situations that challenge that belief. We argue that this state is not emotionally neutral: indeed, we described the act of forming a belief as a relief with respect to the state of doubt. But after the relief passed, it would be consistent with literature on emotional cognition to say that maintaining a certain belief would not be just a peaceful circumstance (Moore & Oaksford, 2002), especially if the agents are vaguely aware that they have, in a way, downplayed their ignorance to feel confident in believing something. We call the epistemic feeling agents have when they interact with the possibility of having one of their beliefs actively challenged, or of finding out that they do not know something, *epistemic discomfort*. Of course, epistemic discomfort is phenomenologically akin to doubt. However, and very importantly, doubt implies an acknowledgment of one’s own ignorance; in contrast, epistemic discomfort does not imply such awareness, as the person needs simply to engage with the possibility that their beliefs are false, and not actively entertain with the idea that they are false.

Before discussing how epistemic discomfort may emerge as a consequence of embublement or informational seclusion online, we think it might be useful to provide a brief vignetta to present how epistemic discomfort may also emerge in offline contexts. The scenario will be useful to clarify exactly what we mean with these terms and to specify that we do not think they represent only online-based occurrences. So, let us consider the following example. Anna, who lives in a small, isolated, and staunch religious community, is walking down the street, when a man, Bob, comes to her to ask for directions. While talking back to him, Anna notices some characteristics about Bob: he’s wearing some eyeliner and nail polish; his midriff is slightly exposed; and his voice is slightly high-pitched. Anna possesses a somewhat traditional conception of masculinity, and has never met an openly homosexual man beforehand. In this scenario, she might feel uncomfortable about her brief encounter with Bob, because she met a man whose apparent characteristics do not conform with her conception of masculinity—so, *what she thinks she knows* about what men are and how they behave. This encounter implicitly (and unwillingly) challenges her assumptions about masculinity because Bob does not conform to that conception; however, that does not mean that she *doubts* the validity of her conception of masculinity. She does not know that her conception of masculinity does not account for instances that contradict it. In other words, she is ignorant about the lack of universal appropriateness of her beliefs about masculinity. However, she does not doubt the validity of her beliefs, because she does not acknowledge that

they might be wrong, and that she might actually be ignorant about what masculinity is or what is supposed to be. The experience of this unacknowledged contradiction causes *not just irritation*, but *discomfort*, because it hints at the possibility that her beliefs are inadequate to account for the designated state of affairs.<sup>20</sup>

So, in case of epistemic discomfort, which are its practical consequences? Anna's encounter with Bob, while implicitly clashing with her beliefs, does not necessarily provide sufficient grounds or motivation for her actually questioning the validity of those beliefs. Instead of questioning the validity of her beliefs about masculinity, she might downplay or deemphasize Bob as an exceptional or aberrant case. Of course, in other circumstances, epistemic discomfort might cause the formation of a doubt; but it might, in others, even strengthen the agent's (inadequate) beliefs instead of pushing her to question their validity. In those cases, epistemic discomfort might so reinforce the agents' ignorance and epistemic bubbles instead of pushing her to change them.

In the next section, we are going to analyze some concrete settings where, we suggest, this dynamic actually occurs: social media. We will argue that the filter bubble should not be understood as a purely algorithmically driven mechanism, but as the result of epistemic discomfort, caused by the way users of social media interact with the beliefs and moral assessments of other people. The reason for this, we maintain, is the way such beliefs and assessments are conveyed to social media users through the platform.

#### 4 Rethinking Filter Bubbles and Echo Chambers: Phenomenology of Belief and Social Feedback on the Web

We will now proceed to reformulate the notion of filter bubble and echo chambers in light of the bubble theses we examined so far. Specifically, we will argue that the filter bubble and echo chambers should be understood as confluences of the epistemic, moral, and ignorance bubbles of Internet users with the way information is configured within the digital platforms they use. The filter bubble and online echo chambers, we argue, are not exclusively generated by recommendation systems, but by the way information and other users' beliefs are presented to one another by the platforms. In other words, it does not matter *what* is presenting the information to people online (i.e., the preference algorithms), but rather the way, or *how* such information is presented to and experienced by the people that interact with it online. The filter bubble and online echo chambers, in other words, are not phenomena purely related to algorithms and what information they present, but to how people react to and interact with information that algorithms present them with—it is not about the information in itself but to people's relation to the information presented

---

<sup>20</sup> With this example we are not in any way implying that this kind of bias or prejudice exists in all religious communities. We do not intend to insult or offend any member of these communities. Rather, this is a mere example, whose purpose is to provide a relatively familiar (to the average reader, at least) instance of a situation where epistemic discomfort is both recognisable and a key factor in the formation of beliefs and their eventual revision. We thank an anonymous peer-reviewer for highlighting the need for this key clarification.

on the platform. When the users encounter information discordant with their previous (embubbled) beliefs, it can cause, we argue, epistemic discomfort, which will tend to push users to stick to their own beliefs and reinforce their epistemic bubble. This is especially true in social networking sites (henceforth SNSs) such as Facebook or Twitter, due to their focus on interpersonal relationships (Cheung & Lee, 2010), which enables people holding contrasting beliefs to interact with opposite standpoints in a *non-mediated* way (as we will discuss further later on). For this reason, we are going to primarily focus our analysis on SNSs.

The argument will proceed as follows. Firstly, we are going to provide a perspicuous representation of digital platforms as environments constituted by information. We will note that even other people are presented to others as sums of the information they provide to the platforms and the content they generate, and that the user's interface of SNSs stands in tension between privacy (as the interface is their own) and public (as other people and sources can access it and interact with it). Secondly, we will build on this characterization of online environments to argue that people might tend to conform and seek standpoints confirming their own opinion due to the unmediated and sometimes unexpected way contrasting standpoints interact. In sum, we argue that epistemic bubbles have a different impact online and offline: in SNSs, the unmediated nature of the interaction fosters epistemic discomfort instead of favoring the emergence of doubt, making people more rigid when they encounter contrasting opinions. We will then support our argument by re-interpreting ambiguous results of studies claiming to have found evidence of filter bubbles and echo chambers, and we will argue that online embubbling, as well as the creation of echo chambers, should be understood as *deliberate*, i.e., as a result of the choices of users, and not purely caused by recommendation systems.

#### 4.1 The Materiality of Others and of Information on the Web

Understandably, SNSs are of special interest for philosophy of mind and epistemology because, from the user's perspective, "there is no gap between information and matter" (Arfini et al. 2019, 382). Digital environments are performed by human and artificial agents<sup>21</sup> that provide users with information, which is primarily (but not exclusively) about people, institutions, and events that exist in the offline world. What Internet users see, however, is just information: while offline in many cases information is situated (socially, culturally, normatively—Cobb, 2001) and mediated by various elements of the context in which people interact, in online platforms users *directly* engage with information without the mediation of a shared and embedding context. While philosophical approaches to cognition that take into account the influence of the environment are not new (see Bateson, 2000; Clark, 2008; Malafouris, 2013; Hutchins, 2010, 1995; Menary, 2010), web-based platforms represent

---

<sup>21</sup> We do not think it relevant to discuss whether the (black-boxed) algorithms that structure and provide information in online platforms can be considered intelligent agents. Our view is compatible with the idea that algorithms simply work as an infrastructure that enables users' information seeking and processing behavior. They do, however, provide the information users consume and interact with: this action may or may not be considered intelligent.

an exceptional instance of cognitive ecology (Smart et al., 2017a, 2017b). Not only are they of exceptional interest because they are constituted by information, but they also afford a very wide range of activities, forms of interaction with information and other users, and allow for varying degrees of personalization.

As Bertolotti et al. (2018) point out, Internet users themselves (and, from an individual user's view, other people online) are configured as bundles of information. Building up on Waite and Bourke (2015), they argue that the association of users of digital platforms with the platform—especially in the case of SNSs—constitutes a case of “cyborgification”: the human user and the used piece of technology are coupled in such a way that the artifact co-constitutes the way the user cognizes and experiences the world (Clark, 2003; Verbeek, 2008, 2011). While the use of an artifact that grants Internet access does not make an agent a cyborg, the same does not hold when we consider the user from within the platform. At the level of abstraction of SNSs—that is, when considering users as part of the interface of SNSs—users are inseparable from the content they produce and the information they provide. The profile settings and information they set up, their network of “friends” online, the information sources they follow and the content they generate—their opinions and expressed values—are what constitutes a user from the perspective of the platform, both in the ways they can act in the platform and in the way they are perceived by others on the platform.

This characterization of digital environments is meant to highlight what might be called the “merely” informational quality of digital platforms. To interact not just with digital platforms, but even with other people on the platforms, means to interact with information—which is structured by the platform's algorithms and provided by other people and agents on the website. While this kind of consideration is not particularly novel, what is relatively unexplored is the way this “merely informational” status of other people, institutions, and information sources affects users' experience of those sources. Due to the meshing of agents and the information they generate on these platforms, one can say that, online, we see other people's beliefs *before* the people themselves. A platform such as Facebook is designed so that its users define themselves based on their beliefs, opinions, desires, preferences, and values—ultimately, based on the content they generate, the sources they follow and the traces they leave on the platform.

Furthermore, the consideration of digital environments as merely informational helps highlight a fundamental ambiguity of many web-based platforms (and, most prominently, social media such as Facebook and Twitter). This ambiguity can be defined as the blurring between private and public space. Consider the case of Facebook's newsfeed, “the constantly updating list of stories in the middle of your home page [which] includes status updates, photos, videos, links, app activity and likes from people, pages and groups that you follow on Facebook” (Facebook, 2021). Through the newsfeed, the information on the platform is perceived and understood from the user's perspective as her own. However, while the experience and structure of these platforms is user-oriented and personalizable, whatever content one produces and consumes also exists in an interpersonal dimension: in sum, from the user's perspective, other agents “fade” into the background of the platform. They

are perceived as part of the informational environment of these websites, and they become objectified and undifferentiated as bits of information.

It is both the natural limitations of computer-mediated communication and the design choices of these platforms (the constitution of a profile and how it explicitly invites you to post what is on your mind) that lead users to consider not only each other, but eventually also themselves, as their generated content. “Online, it is often easier to separate people from their embodied experiences, or to mistake the part for the whole—or to never even see the whole, and therefore never understand the context from which a particular collection of pixels has been unmoored” (Phillips and Milner, 2018, 89). The settings and structure of the platform leads to the understanding of other agents online only based on what they post and share, ignoring that they are people grounded in an offline—and unseen—context and merged with a platform that transforms their experience and cognition. Or, as Nelson (2018) puts it: “Social media allows us to persistently emphasize who we are and set aside the question of what we are altogether” (178).

## 4.2 Context Collapse and the Unmediated Experience of Information Online

We now turn to the analysis of what we have called the “immediacy” or “unmediatedness” that characterizes the way users experience information in many different platforms, and mainstream SNSs most prominently. The specificity of this immediacy originates in an issue intrinsic to computer-mediated communication already individuated by Kiesler et al. (1984). They wrote: “[c]ommunicators must imagine their audience, for at a terminal it almost seems as though the computer itself is the audience” (p. 1125). One of the most trivial differences between online and face-to-face interaction is the fact that the computer user does not *see* the recipient of her communicative acts. While through media such as email or private instant messages the perception of the message recipient can be relieved by the knowledge of who is going to consume that information (i.e., you know who you are messaging) the problem of imagining who is going to consume the information you produce is prominent in platforms such as SNSs. For in these platforms, the content users consume, share, and produce is potentially shared to and by the entirety of their online social network, which generally includes people from very different offline social contexts—your friends, your co-workers, your family members, and mere acquaintances to name a few. Even strangers, under certain circumstances, can come into contact with your posts, comments, and reactions online. This feature of many contemporary forms of computer-mediated interaction and communication is known as context collapse: “[T]he flattening out of multiple distinct audiences in one’s social network, such that people from different contexts become part of a singular group of message recipients” (Vitak, 2012, p. 451).

The concept gained much attention since its introduction by danah boyd (2008; Marwick and boyd, 2011; Gil-Lopez et al., 2018) to describe how the indeterminacy of the audience on SNSs such as Facebook and Twitter leads users to moderate content production and self-presentation. Because people do not really know who is going to actually consume the content and how would they react, many Twitter users feel the need to imagine what is appropriate to post to match the expectations of all possible content consumers—friends, family members, strangers, and so on. However, while

context collapse has been generally studied in relation to privacy management (Vitak, 2012; Marwick and boyd, 2014) and exposure to news online (Beam et al., 2018; Kim & Ihm, 2020), there is still some uncertainty in the literature regarding its exact conceptualization and implications (Davis & Jurgenson, 2014; Costa, 2018; Szabla and Bloemmert, 2020).

For the purposes of this paper, context collapse is relevant because it constitutes the source of what we called the *immediacy* or unmediatedness of information online. Specifically, the immediacy of information is determined by context collapse because the merging of different social groups leads to the indetermination of normative standards for communication and interaction with others. The impact of context collapse in online self-presentation is due to the fact that different (offline) social groups have different expectations for what it is appropriate to say or not to say. The various expectations, more or less acknowledged standards for appropriate and inappropriate behavior that underlie social interaction in different face-to-face settings become undetermined. The audience of the content you produce sets normative standards for how you choose to behave and what you say. While you would behave differently in front of strangers and with friends or family members, online all of these audiences, as consumers of the content you produce, are reduced to one. Online, expectations regarding the consumption and origin of content, as well as interaction with others, becomes indeterminate. The information consumed on SNSs is unmediated because there are not universally shared norms for interaction that can help framing and interpret that information univocally—a lack of frame that can cause an experience of that information as unexpected, unmotivated, or even unjustified.

### 4.3 Reinforcing Epistemic, Moral, and Ignorance Bubble online

So, let us refocus on the impact of the bubble theses if we take into consideration the merely informational nature of digital platforms from the phenomenological perspective of users. Let us consider the following scenario. Andy and Betty are two Facebook users who are on the platform to consume the news. Andy firmly believes that the vaccine against COVID-19 is safe and effective: he is aware of the wide amount of misinformation and conspiracy theories out there, and he firmly believes that scientific research cannot (epistemic belief) and should not (moral belief) be disputed from non-scientific perspectives, such as politicians or religious institutions. Betty, on the other hand, is not as sure about the safety and effectiveness of the vaccine: she knows that online there is much contradictory information on the matter, and she does not necessarily have a strong opinion regarding which sources have a higher degree of authority on the matter. Suppose that, while on Facebook, they independently stumble on the same post: for instance, a post with a hyperlinked article that mentions how a religious institution deems the Johnson & Johnson vaccine dangerous (epistemic information) and immoral (moral information). Andy swiftly writes, in the comment section of the post, that in no way a religious institution is entitled to deem the validity (epistemic bubble) or morality of a scientific practice (moral bubble), and that he's appalled by people who ignore or do not just follow the scientific consensus

in this regard. Betty, in contrast, reads not just the post, but also Andy's comment. She feels uncomfortable and offended by this comment: she does not "just follow" the scientific consensus, as her feelings and beliefs about the vaccine also come from her social group and, in general, sources that she deems reliable (epistemic bubble) and just (moral bubble), which Andy does not consider. She therefore comments as well: she calls out Andy for his ignorance regarding other people's sources and questions his proclaimed entitlement regarding who can express her opinion about what. After this brief interaction, Andy will feel even more sure about his outstanding opinion regarding vaccines, as his experience of the alternative view was still negative (reinforcing both his epistemic and moral bubbles); while Betty will feel even more insecure regarding vaccines, as she experienced the interaction with alternative view(s), as violent and wrong (reinforcing her bubbles, as well).

This short story regarding a non-constructive interaction between two social media users is a paradigmatic case that stitches together all the various conceptual threads we presented in the course of the paper. In particular, this interaction constitutes an example of epistemic discomfort in online settings: if we take the perspective of either Andy or Betty, what we see is an agent who is confronted with someone else's perspective that questions the veracity of her epistemic and moral bubble. The "evidence" the agent experiences challenges her and pushes her to consider her ignorance on the matter at hand—here, on the safety and morality of vaccines. The agent, here, does not experience doubt: s/he does not acknowledge the possibility that s/he might, in fact, be ignorant and wrong on the matter at hand. In a sense, there are not even sufficient grounds for his/her to do so: s/he is just told by another person that she is wrong about it.<sup>22</sup> In this case, epistemic discomfort does not lead to a reconsideration of the agent's views, but to their confirmation—to a reinforcement of his/her epistemic and moral bubbles.

This dynamic is enabled by the specific conditions of online environments. As we have pointed out, in digital platforms everything the Internet user sees is information, even in the case of other people. While Betty cannot see or know about Andy's frustration regarding misinformation on COVID-19 online, Andy cannot know or predict how Betty (or other people) consider a source reliable and their motivations to do so. Andy only sees a Facebook post that he considers (epistemically and morally) wrong, and is appalled that other people do not agree with him; and Betty only sees someone who criticizes anyone who disagrees with his ideas. The contrary perspective is experienced as unmediated, decontextualized and unjustified in tone and scope; the way it is encountered—which, it must be noted, arises from the way the platform is designed and configured—causes epistemic discomfort that is not yet sufficient for the explicit

---

<sup>22</sup> It is noteworthy that, in line with the moral bubble thesis, the agent's cognitive rigidity in front of epistemic discomfort is also motivated by the fact that the belief in question may identify her as a member of her (moral and epistemic) community. Epistemic and (especially) moral embubbling can be shared by a community. For an agent to acknowledge and explicate doubt in this situation would entail that she does not affiliate with her group and its shared values, which constitutes further reasons for her sticking to her own epistemic and moral bubble (see Kahan et al. 2012).



formulation of doubt. Therefore, Betty (or Andy) can choose to discredit the validity of the contrary perspective (or, as Nguyen [2020] puts it, epistemically discrediting it).

The confrontation between agents or information sources with contrasting perspectives and opinions leads them to perceive unmediated clashes between opposite and almost incompatible views (so epistemic and moral bubbles). The immediacy of this clash can lead to the reinforcement of the epistemic bubble instead than to an open reflection of one's own (possible) ignorance. This reinforcement, if offline would not have been sufficiently motivated, if considered within online environments should not be necessarily described as irrational. Due to context collapse, the merging of different expectations for what is appropriate to say, share, or behave that are grounded in different offline situations contributes to a generalized uncertainty regarding how to react in front of unexpected situations. As Rini (2017) points out, on the contrary of face-to-face interaction, norms of communication on social media are disputed: the aims, purposes, and causes of a specific post are not undisputedly understood by social media users. Users can only assess other people by what they actually post; they lack both access to other users' motivations and background, as well as a shared understanding of norms of communication and behavior. In this context, taking an epistemically partisan stance toward contrary opinions—that is, reinforcing one's own previous belief in front of contrary standpoints—can be considered as a rational choice. Because the opinions and claims of other people online come to be in an underdetermined context, sticking to one's own perspective can provide interpretative support of information and claims generated from others. It does so by implicitly establishing who is an epistemic and moral peer and who is not and by elucidating one's own values. When one experiences epistemic discomfort online by encountering unmediated and seemingly unjustified opinions (a feeling that may be expressed with the words: "How can you say that?") reinforcing one's own epistemic, ignorance, and moral bubbles can be seen as the most intuitive strategy.<sup>23</sup>

#### 4.4 Redefining Filter Bubbles and the Arbitrariness of Epistemic Isolation

We now have all the resources to re-formulate the notion of filter bubble by changing the focus of the original problem and taking into account the empirical evidence available regarding information exposure online. Pariser understood the filter bubble as a form of technologically caused informational and epistemic seclusion, where contrary perspectives are excluded by the preference algorithms. In contrast, we reflected on the phenomenology of belief and of experiencing contrary perspectives, and we looked at the transformation of this experience in online environments. We proposed that the experience of epistemic discomfort in web-based platforms is especially due to the way others' opinions are framed and presented. The

---

<sup>23</sup> This reinforcement of epistemic, ignorance and moral bubble can also be understood as rational (all things considered) if we interpret it as a sort of adaptive preference mechanism. This notion is used to describe the formation of an opinion in front of external pressure or of a reduced number of options (Elster 1983).

unmediated manner other opinions are experienced and how easily they can cause epistemic discomfort in digital environments more reasonably lead to sticking to and maintaining one's own outstanding opinion than to consciously doubting it, because in the user's perspective there are not sufficient grounds for doubt.

We therefore define the filter bubble as *the reinforcement of one's epistemic, moral, and ignorance bubble caused by epistemic discomfort experienced online*. We do not see the filter bubble as caused purely by technology, but by the interaction between user and technology, and specifically by the way users experience each other's views in online settings. The unmediated way contrary perspectives are encountered thanks to context collapse lead to an uncomfortable experience, as the user's outstanding belief is questioned. However, due to the indeterminacy and ambiguity of the conditions surrounding the contrary opinion, users will more probably feel prompted to stick to their outstanding opinion instead of doubting it: when the options are either to assume one's own opinion is correct or to side with a standpoint that questions her own without apparent grounds, that can even be considered reasonable, despite its being irrational.

This technologically mediated experience of epistemic discomfort leads, we suspect, to a transformation in attitude toward one's own beliefs and the perspective of others. Your outstanding belief, which is being challenged, instead of being the object of your evaluation, it might assume a guiding function toward your evaluation of others' perspectives, and end up taking a normative role. Because of the uncertainty regarding interpreting and evaluating information online as valid, authoritative or justified, when experiencing epistemic discomfort one can rationally solve it by appealing to one's outstanding belief as valid—even though it might not be necessarily so. Exposure to contrary or diverse perspectives in these settings can lead not to open-mindedness, but to reinforcing one's own both epistemic and moral self-righteousness—because one's own outstanding beliefs become the only reference points while navigating online environments. In this climate of informational uncertainty, to follow one's own (presumed) knowledge can be considered the most rational—and definitely the least unpleasant—course of action. Unmediated, and eventually unintentional, exposure to contrary perspective would not lead to open-mindedness, but to what Nguyen (2020) calls a “disagreement-reinforcement mechanism” (147): the existence of contrary perspectives can lead to the perceived corroboration of one's own outstanding belief, especially when the believer expects that other people will attempt to contradict her.<sup>24</sup> One's own epistemic bubble is, in this scenario, what supports the filtering of information and the reference point for evaluating other data and others' perspectives. In this sense, one might say that the filter bubble is “already in your head”—your epistemic bubble will be reinforced by the experience of epistemic discomfort online and not just remain unquestioned, but drive information consumption behavior.

---

<sup>24</sup> See also Begby (2020) for a thorough analysis of this confirmation mechanism through confrontation with contrary perspectives.

At this point, it must be noted that we are not arguing that all Internet users are going to experience epistemic discomfort in the same way. It is a truism to say that some people are more open-minded than others (and that would probably remain true also in online environments). However, we assume that it is specifically the way most social media users feel challenged online that prompts them to conform to and reinforce their outstanding beliefs. In other words, exposure to contrary perspectives—in a decontextualized, unmediated, and distorted manner typical of most web-based platforms—makes one's own epistemic bubbles to be more rigid, and to filter out contradicting opinions. Not informational seclusion, we argue, but the way individual users manage unmediated contrasting opinions leads to the cognitive rigidity and tendential polarization that Pariser and Sunstein worry about.

We now turn to test our argument by looking at a group of studies that found some empirical evidence of algorithmically caused informational seclusion, either in the form of filter bubbles or of echo chambers. As Bruns (2019) points out, we should remember that not only the filter bubble is generally understood as algorithmically generated, but even echo chambers are sometimes seen as indirectly caused by preference algorithms. We intend to re-interpret the findings of these studies taking into account also the way users feel about (contrary) information they find online.

Flaxman et al. (2016) analyzed the behavior of more than 50,000 Internet users' news consumption on 100 online news platforms. They were specifically researching whether the consumption of news provided via social media and web search would have been associated with a higher degree of ideological segregation (of a lesser degree of contact with contrary information) than directly visiting a news website's page. They did find that social media and web search was indeed (slightly) associated with ideological segregation and information consumption of like-minded news. However, the ideological segregation they found generally reflected offline patterns of news consumption and social media and web search were also associated with *exposure* to a higher variety of outlets than direct search did. In other words, while people are more exposed to cross-cutting news outlets when using social media and web search (in contrast to visiting the outlet's web page directly), they seem to engage with such cross-cutting outlets less. As the authors put it: "[A]rticles found via social media or web-search engines are indeed associated with higher ideological segregation than those an individual reads by directly visiting news sites. However, we also found, somewhat counterintuitively, that these channels are associated with greater exposure to opposing perspectives" (318).

This finding seems to question the notion of filter bubble as an algorithmically generated informational seclusion. After all, users of web search engines and social media are exposed to contrary perspectives, contradicting Pariser and Sunstein; however, they do not engage with those outlets despite the fact that they are exposed to them. Our account of filter bubble, which shifts the attention from the "mere" platform to the user's cognitive processes, can explain Flaxman et al.'s findings as follows: incidental exposure to contradictory perspectives, which will be generally perceived as groundless, causes epistemic discomfort, a reinforcement of one's own outstanding beliefs, and, ultimately, to the inclination to consume like-minded

information. It is, in other words, in virtue of unmediated exposure to contrary perspectives that Web users become ideologically segregated.

The same dynamic of reinforcement of one's own outstanding beliefs in face of unmediated contrary information is also helpful in explaining the formation of echo chambers. As we saw in Section 2.3, one of the problems with the studies that claimed to have found evidence for echo chambers (Quattrociocchi et al., 2016; Del Vicario et al., 2016; Wollabæk et al., 2019) is the presumption that the online community that constitutes it is very much isolated from contrary information sources. This is actually a pretty heavy assumption: a higher degree of activity with like-minded individuals does not entail that users do not get in contact *at all* with contrary perspectives—at least not on mainstream websites such as Facebook and Twitter. Therefore, one could interpret the notion of echo chamber in a “softer” way, as a quantitative difference in interaction with like-minded vs. contrary users and information sources. This “softer” interpretation of the notion, however, needs to still explain the effect of the outgroup's opinions and perspectives on individual members of the echo chamber, in light of the fact that digital platforms users are exposed to them.

Our understanding of the filter bubble, as the result of and manner of avoiding epistemic discomfort in online settings, can account for the formation of echo chambers and their maintenance in spite of exposure to contrary perspectives. The experience of unmediated opinions that contradict our own can lead to reinforcing our outstanding beliefs and, ultimately, to cognitive rigidity in front of contrary information. And while research that investigates echo chambers agrees that their members are driven to conformity by a confirmation bias, the previous experience and (mis)management of epistemic discomfort, as well as the desire to avoid it in future digital interaction explain why such bias would develop in the first place—even in agents who were not polarized to begin with. The desire to avoid epistemic discomfort and filter out contrary information can then lead to conformity to whatever information source will confirm and conform to our outstanding belief, including other people who share our beliefs, views and attitudes. In his original essay (2005) on the notion of epistemic bubble, Woods noted that the epistemically closer our source of feedback is, the least can our epistemic bubble burst. In online environments, where information from diverse sources appears qualitatively undistinguishable, seeking interaction with like-minded individuals can reinforce one's own epistemic bubble and drive information evaluation and consumption.

Our re-elaboration of the filter bubble has a further implication, which stands in stark contrast with the central worries that Pariser and Sunstein underlined. The generalized sentiment that pervades their understanding of filter bubbles and echo chambers consists in the fear that digital technologies will cause intellectual isolation and, therefore, polarization. Unhealthy information consumption and hostility toward contrary perspectives would be explained by the fact that the platform does not allow users to get in contact with different perspectives—a dynamic that empirical evidence that does not presuppose the existence of echo chambers seems to substantially prove wrong. In contrast, our re-formulation of the notion of filter bubble, by taking a step back and looking at how a user feels in the interaction with web-generated information, understands the influence of digital technologies in quite a

different way. While it is true that online platforms such as social media enable a high degree of personalization in information management, we argue that it is the modality of experience of contrary perspectives online that causes the reinforcement of one's outstanding beliefs. It is the unmediated way opposite viewpoints are experienced that leads to conformity and to the discrimination of those perspectives. In this sense, exposure to contrary perspectives in online settings is actually functional to the creation of filter bubbles and echo chambers. The reason being that such exposure is unmediated and often perceived as "out of the blue," as groundless and unjustified: encountering opposing views in online settings causes epistemic discomfort that leads to rigidity more than open-mindedness. The formation of echo chambers and the individual development of cognitive rigidity is, in other words, the result of users' choice, not of the algorithm's decision.

## 5 Concluding Thoughts

In contrast to previous conceptions of the influence of digital technologies on everyday epistemic practices, our interpretation of the filter bubble has some important theoretical and practical advantages. Firstly, describing the filter bubble as the result of user deliberation (encouraged by digital design) caused by epistemic discomfort does not rely on any form of technological determinism.<sup>25</sup> Instead of attributing the formation of polarized (and potentially secluded) users and communities on black-boxed algorithmic systems, we examine the relationship between the information presented on digital platforms and the belief-formation and epistemic feelings of the user. This shift of focus distributes the responsibility for epistemic rigidity and polarization in between all parties involved in a more nuanced manner. Secondly, our elaboration of the filter bubble as the primary cause of polarization and rigidity is coherent with already existing empirical research that investigates information exposure and consumption behavior, much more so than the original filter bubble thesis and the idea of algorithmically generated echo chambers do. Thirdly, our argument can account for different motivations behind users' information management online. On the one hand, epistemic discomfort online can also be managed with a higher degree of open-mindedness (even though we take it that that would be the exception more than the norm). On the other hand, the kind of cognitive rigidity that users may feel online and the desire to avoid it is shared by both politically savvy users who use these platforms primarily to seek political information, and those who use them for other reasons and simply stumble upon information. Fourthly, our account identifies a causal relation, instead of a mere correlation, between information consumption behavior in online settings and the formation of a polarized and partisan attitude toward contrary perspectives. And finally,

---

<sup>25</sup> We want to make clear that the bubble theses and our re-conceptualization of the filter bubble do not hinge on any form of determinism (psychological, sociological, metaphysical etc.): while they do describe some epistemological limits of human agency, such limits do not force us to believe certain things instead of others, or make our belief-formation process as such entirely predictable—i.e. these limits are not binding in the same way some *laws* would be. We thank an anonymous peer reviewer for making explicit to us the relevance of this clarification.

our account can be put into test through empirical research, as the experience of epistemic discomfort and the way it is managed can be explored among social media users in general through qualitative research.

Of course, still relevant questions and issues pertain to the idea of a filter bubble as so reframed. One problem over all is practical: if users in digital environments are prone to reinforce their extremely personal and cognitively relevant epistemic, ignorance, and moral bubbles, finding a way to break this mechanism would involve both changes in the interface's design and in the ways users will interact with it. On the one hand, the implementation of media literacy policies that can support users' understanding of how these technologies work and of their own situatedness (and intrinsically limited perspective) within these platforms would be more than adequate—even though such media literacy might be radically different than for traditional media (Phillips & Milner, 2021). On the other hand, however, we believe that the most effective way of tackling the problem of filter bubbles, intended as epistemic discomfort experienced by people on digital platforms, would be to implement some design changes to the interface of these platforms. Specifically, because context collapse and the interpretative uncertainties it implies are one of the primary factors contributing to epistemic discomfort, through the implementation of designed norms it would be possible to contextualize the information people run into online. Such norms may not only help users keep in mind that their personal expectations might be inadequate for the content they encounter, but actually support a more open mind and higher degree of understanding toward others—not only helping frame content that appears groundless, but actively encouraging a more constructive attitude toward the information they encounter.

That being said, while we do think that changes in the design's interface that tackle context collapse would be the most efficient solution, we do not think of it as the only possible one. We are in good support of a plurality of strategies for making people more intellectually humile, or more open-minded—including not only changes in the interface, but also media literacy policies with that purpose. Eventual changes in the interface's design to reduce epistemic discomfort will also need to be not conflicting too strongly with the service provider's corporate interests. Thus, solving this specific problem will probably require the implementation of socio-technological ways to encourage a more constructive epistemic interaction on digital platforms. Given the complexity of this follow-up problem, we look forward to actively exploring its implications in future research.

**Acknowledgements** We would like to thank Marc V.P. Slors and Antonia Leise for their feedback and intellectual support in the making of this paper. We also thank two anonymous peer-reviewers, whose feedback greatly improved this paper on the way to publication.

**Authors' contributions** Giacomo Figà Talamanca wrote sections 1, 2 and 4 of the manuscript, while Selene Arfini wrote section 3. All authors contributed to manuscript revision, read and approved the submitted version. Both authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Funding** Giacomo Figà Talamanca's contribution is supported by a scholarship by Radboud University for Honors Academy Project "Rethinking Filter Bubbles." Selene Arfini's contribution is supported by University of Pavia and PRIN 2017 Research 20173YP4N3—MIUR, Ministry of University and Research, Rome, Italy.

## Declarations

**Data Availability** Not applicable.

**Code availability** Not applicable.

**Ethical Approval** Not applicable.

**Consent to Participate** Not applicable.

**Consent to Publish** Not applicable.

**Conflicts of interest/Competing interests** The authors declare no competing or conflicting interests.

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

## References

- Arango-Muñoz, S. (2014). The Nature of Epistemic Feelings. *Philosophical Psychology*, 27(2), 193–211. <https://doi.org/10.1080/09515089.2012.732002>
- Ackerman, P., Beier, M., & Bowen, K. (2002). What we really know about our abilities and our knowledge. *Personality and Individual Differences*, 33, 587–605. [https://doi.org/10.1016/S0191-8869\(01\)00174-X](https://doi.org/10.1016/S0191-8869(01)00174-X)
- Arfini, S., (2019). *Ignorant Cognition A Philosophical Investigation of the Cognitive Features of Not-Knowing*. Springer Verlag
- Arfini, S., Bertolotti, T., & Magnani, L. (2019). Online communities as virtual cognitive niches. *Synthese*, 196, 377–397. <https://doi.org/10.1007/s11229-017-1482-0>
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to Ideologically Diverse News and Opinion on Facebook. *Science*, 348(6239), 1130–1132. <https://doi.org/10.1126/science.aaa1160>
- Bateson, G. (2000). *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. University of Chicago Press.
- Baber, C., Chemero, T., & Hall, J. (2019). What the Jeweller's Hand Tells the Jeweller's Brain: Tool Use, Creativity and Embodied Cognition. *Philosophy & Technology*, 32, 283–302. <https://doi.org/10.1007/s13347-017-0292-0>
- Beam, M. A., Child, J. T., Hutchens, M. J., & Hmielowski, J. D. (2018). Context Collapse and Privacy Management: Diversity in Facebook Friends Increases Online News Reading and Sharing. *New Media & Society*, 20(7), 2296–2314. <https://doi.org/10.1177/1461444817714790>
- Beam, M. A., & Kosicki, G. M. (2014). Personalized News Portals: Filtering Systems and Increased News Exposure. *Journalism & Mass Communication Quarterly*, 91(1), 59–77. <https://doi.org/10.1177/1077699013514411>
- Begby, E. (2020). Evidential Preemption. *Philosophy and Phenomenological Research* 102: 515-530. <https://doi.org/10.1111/phpr.12654>
- Bertolotti, T., Arfini, S., & Magnani, L. (2018). Cyber-Bullies as Cyborg-Bullies. In R. Luppincini (Ed.), *The Changing Scope of Technoethics in Contemporary Society* (pp. 60-73). IGI Global. <https://doi.org/10.4018/978-1-5225-5094-5.ch004>



- Borra, E., and Weber, I. (2012). Political Insights: Exploring Partisanship in Web Search Queries. *First Monday* 17(7). <https://doi.org/10.5210/fm.v17i7.4070>.
- boyd, d., 2008. *Taken Out of Context: American Teen Sociality in Networked Publics*. PhD Dissertation. University of California-Berkeley, School of Information.
- Boyd, R., & Richerson, P. J., (2001). Norms and Bounded Rationality. In Gigerenzer, G. & R. Selten (Eds.) *Bounded Rationality. The Adaptive Toolbox* (pp. 281–296). Cambridge: The MIT Press.
- Bruns, A. (2019). *Are Filter Bubbles Real?* John Wiley & Sons.
- Bruns, A., Moon, B., Münch, F., & Sadkowsky, T. (2017). The Australian Twittersphere in 2016: Mapping the Follower/Followee Network. *Social Media + Society* 3(4):1–15. <https://doi.org/10.1177/2056305117748162>.
- Büchi, M., Fosch Villaronga, E., Lutz, C., Tamò-Larrieux, A., Velidi, S., & Viljoen, S. (2019). Chilling Effects of Profiling Activities: Mapping the Issues. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3379275>
- Burr, C., Cristianini, N., & Ladyman, J. (2018). An Analysis of the Interaction Between Intelligent Software Agents and Human Users. *Minds & Machines*, 28, 735–774. <https://doi.org/10.1007/s11023-018-9479-0>
- Cadwalladr, C. (2018). "I Created Steve Bannon's Psychological Warfare Tool": Meet the Data War Whistleblower. *The Guardian*. Retrieved 18 March 2021 ([theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump](https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump)).
- Cardenal, A. S., Aguilar-Paredes, C., Galais, C., & Pérez-Montoro, M. (2019). Digital Technologies and Selective Exposure: How Choice and Filter Bubbles Shape News Media Exposure. *The International Journal of Press/politics*, 24(4), 465–486. <https://doi.org/10.1177/1940161219862988>
- Cassam, Q. (2018). *Vices of the Mind: From the Intellectual to the Political*. Oxford University Press.
- Cheung, C. M. K., & Lee, M. K. O. (2010). A Theoretical Model of Intentional Social Action in Online Social Networks. *Decision Support Systems*, 49(1), 24–30. <https://doi.org/10.1016/j.dss.2009.12.006>
- Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.
- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. OUP USA.
- Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58(1), 7–19. <https://doi.org/10.1093/analys/58.1.7>
- Clarke, M. (1990). Epistemic Norms and Evolutionary Success. *Synthese*, 2, 231–244. <https://doi.org/10.1007/BF00484793>
- Clowes, R. W. (2015). Thinking in the cloud: The cognitive incorporation of cloud-based technology. *Philosophy & Technology*, 28(2), 261–296. <https://doi.org/10.1007/s13347-014-0153-z>
- Cobb, P. (2001). Situated Cognition: Origins. *International Encyclopedia of the Social & Behavioral Sciences* 14126–29.
- Costa, E. (2018). Affordances-in-Practice: An Ethnographic Critique of Social Media Logic and Context Collapse. *New Media & Society*, 20(10), 3641–3656. <https://doi.org/10.1177/1461444818756290>
- Davies, H. C. (2018). Redefining Filter Bubbles as (Escapable) Socio-Technical Recursion. *Sociological Research Online*, 23(3), 637–654. <https://doi.org/10.1177/1360780418763824>
- Davis, J. L., & Jurgenson, N. (2014). Context collapse. Theorizing context collusions and collisions. *Information, Communication and Society*, 17(4), 476–485. <https://doi.org/10.1080/1369118X.2014.888458>
- Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G. & Quattrocioni, W. (2016). Echo Chambers: Emotional Contagion and Group Polarization on Facebook. *Scientific Reports* 6: 37825. <https://doi.org/10.1038/srep37825>
- Dunning, D. (2011). The Dunning-Kruger Effect: On Being Ignorant of One's Own Ignorance. *Advances in Experimental Social Psychology*, 44, 247–296. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Dutton, W. H., Reisdorf, B., Dubois, E., & Blank, G. (2017). Social Shaping of the Politics of Internet Search and Networking: Moving beyond Filter Bubbles, Echo Chambers, and Fake News' *Quello Center Working Paper No. 2944191*. <https://doi.org/10.2139/ssrn.2944191>.
- Elster, J. (1983). *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge University Press.
- Evans, J.S.B.T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology* 59, 255–278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Facebook (2021). How News Feed Works | Facebook Help Center. Retrieved 11 May 2021(<https://www.facebook.com/help/1155510281178725?ref=u2u>).

- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly*, 80(S1), 298–320. <https://doi.org/10.1093/poq/nfw006>
- Fletcher, R., Cornia, A., & Nielsen, R. K. (2020). How Polarized Are Online and Offline News Audiences? A Comparative Analysis of Twelve Countries. *The International Journal of Press/politics*, 25(2), 169–195. <https://doi.org/10.1177/1940161219892768>
- Fletcher, R., & Nielsen, R. K. (2018a). Are People Incidentally Exposed to News on Social Media? A Comparative Analysis. *New Media & Society*, 20(7), 2450–2468. <https://doi.org/10.1177/146144817724170>
- Fletcher, R., & Nielsen, R. K. (2018b). Automated Serendipity: The Effect of Using Search Engines on News Repertoire Balance and Diversity. *Digital Journalism*, 6(8), 976–989. <https://doi.org/10.1080/21670811.2018.1502045>
- Gabriel, T., & Hannan, R. (2011). *Islam and the veil: Theoretical and regional contexts*. Continuum. <https://doi.org/10.5040/9781472548559>
- Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2018). Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship. In *Proceedings of the 2018 World Wide Web Conference* (pp. 913–22).
- Gil-Lopez, T., Shen, C., Benefield, G. A., Palomares, N. A., Kosinski, M., & Stillwell, D. (2018). One Size Fits All: Context Collapse, Self-Presentation Strategies and Language Styles on Facebook. *Journal of Computer-Mediated Communication*, 23(3), 127–145. <https://doi.org/10.1093/jcmc/zmy006>
- Griffiths, P., & Wilkins, J. (2010). When Do Evolutionary Explanations of Belief Debunk Belief? *Darwin in the 21st Century*: 5314.
- Grimmelmann, J. (2014). As Flies to Wanton Boys. The Laboratorium. Available at: [https://laboratorium.net/archive/2014/06/28/as\\_flies\\_to\\_wanton\\_boys](https://laboratorium.net/archive/2014/06/28/as_flies_to_wanton_boys). (Accessed 13th February 2021)
- Gross, M., & McGoey, L. (Eds.). (2015). *Routledge International Handbook of Ignorance Studies*. Routledge.
- Gurumurthy, A., & Bharthur, D. (2018). Democracy and the algorithmic turn. *International Journal on Human Rights* 27:39. <https://sur.conectas.org/en/democracy-and-the-algorithmic-turn/>. Accessed 7 Sep 2021
- Haim, M., Graefe, A., & Brosius, H.-B. (2018). Burst of the Filter Bubble? *Digital Journalism*, 6(3), 330–343. <https://doi.org/10.1080/21670811.2017.1338145>
- Hildebrandt, M., Kerdels, J., Albiez, J., & Kirchner, F. (2008). Robust Vision-Based Semi-Autonomous Underwater Manipulation. *Intelligent Autonomous Systems* 10: IAS-10, 308.
- Holt, S. S., Loucks, N., & Adler, J. R. (2009). Religion, Culture, and Killing. Pp. 1–6 In Loucks, N., S. S. Holt, & J. R. Adler (Eds.). *Why We Kill. Understanding Violence across Cultures and Disciplines*, pp. 1-6. London: Middlesex University Press.
- Hookway, C. (1998). Doubt: Affective States and the Regulation of Inquiry\*. *Canadian Journal of Philosophy Supplementary*, 24, 203–225. <https://doi.org/10.1080/00455091.1998.10717500>
- Hutchins, E. (1995). *Cognition in the Wild*. MIT press.
- Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2(4), 705–715. <https://doi.org/10.1111/j.1756-8765.2010.01089.x>
- Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Indiana University Press.
- Ihde, D., & Malafouris, L. (2019). *Homo faber* Revisited: Postphenomenology and Material Engagement Theory. *Philosophy & Technology*, 32, 195–214. <https://doi.org/10.1007/s13347-018-0321-7>
- Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The Polarizing Impact of Science Literacy and Numeracy on Perceived Climate Change Risks. *Nature Climate Change*, 2(10), 732–735. <https://doi.org/10.1038/nclimate1547>
- Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, 39(10), 1123–1134. <https://doi.org/10.1037/0003-066X.39.10.1123>
- Kim, E.-M., & Ihm, J. (2020). Online News Sharing in the Face of Mixed Audiences: Context Collapse, Homophily, and Types of Social Media. *Journal of Broadcasting & Electronic Media*, 64(5), 756–776. <https://doi.org/10.1080/08838151.2020.1835429>
- Krafft, T.D., Gamer, M. & Zweig, K.A. (2019). What did you see? A study to measure personalization in Google’s search engine. *EPJ Data Science* 8:38. <https://doi.org/10.1140/epjds/s13688-019-0217-5>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6), 1121. <https://doi.org/10.1037/0022-3514.77.6.1121>.

- Kruger, J., & Dunning, D. (2002). Unskilled and unaware—but why? A reply to Krueger and Mueller. *Journal of Personality and Social Psychology*, 82(2), 189–192. <https://doi.org/10.1037/0022-3514.82.2.189>
- Lapowsky, I. (2019). ‘How Cambridge Analytica Sparked the Great Privacy Awakening. Repercussions from the scandal swirling around the data analytics firm continue to be felt across the tech industry.’ *Wired*. Retrieved 18 March 2021 (<https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/>)
- Magnani, L. (2011). *Understanding Violence: The Intertwining of Morality, Religion and Violence: A Philosophical Stance*. Springer-Verlag.
- Malafouris, L. (2013). *How Things Shape the Mind*. MIT press.
- Marciano, A., Nicita, A., & Ramello, G. B. (2020). Big data and big techs: Understanding the value of information in platform capitalism. *European Journal of Law and Economics*, 50, 345–358. <https://doi.org/10.1007/s10657-020-09675-1>
- Marwick, A. E., & boyd, d. (2011). I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society*, 13(1), 114–133. <https://doi.org/10.1177/1461444810365313>
- Marwick, A. E., & boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, 16(7), 1051–1067. <https://doi.org/10.1177/1461444814543995>
- McKay, R. T., & Dennett, D. C. (2009). The Evolution of Misbelief. *Behavioral and Brain Sciences*, 32, 493–510.
- Menary, R. (2010). *The Extended Mind*. Mit Press.
- Moore, S. C., & Oaksford, M. (Eds.). (2002). *Emotional Cognition: From Brain to Behaviour*. John Benjamins Publishing.
- Nechushtai, E., & Lewis, S. C. (2019). What Kind of News Gatekeepers Do We Want Machines to Be? Filter Bubbles, Fragmentation, and the Normative Dimensions of Algorithmic Recommendations. *Computers in Human Behavior*, 90, 298–307. <https://doi.org/10.1016/j.chb.2018.07.043>
- Nguyen, C. T. (2020). Echo Chambers and Epistemic Bubbles. *Episteme*, 17(2), 141–161. <https://doi.org/10.1017/epi.2018.32>
- Nuhfer, E., Fleisher, S., Cogan, C., Wirth, K., & Gaze, E. (2017). How Random Noise and a Graphical Convention Subverted Behavioral Scientists’ Explanations of Self-Assessment Data: Numeracy Underlies Better Alternatives. *Numeracy*, 10(1), 4. <https://doi.org/10.5038/1936-4660.10.1.4>
- Ochigame, R., & Ye, K. (2021). Search Atlas: Visualizing Divergent Search Results Across Geopolitical Borders. *Designing Interactive Systems Conference 2021 (DIS '21)*: 1970–1983. <https://doi.org/10.1145/3461778.3462032>
- O’Reilly, T. (2007) What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *Communications & Strategies*, 1 17. Available at SSRN: <https://ssrn.com/abstract=1008839>
- Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.
- Peels, R., & Blaauw, M. (2016). *The Epistemic Dimensions of Ignorance*. Cambridge University Press.
- Peirce, C. S. (2011). The Fixation of Belief. In Talisse, R.B., and S.F. Aikin (Eds.) *The Pragmatist Reader*, pp. 37–49. Princeton: Princeton University Press.
- Phillips, W., & Milner, R. M. (2018). *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online*. John Wiley & Sons.
- Phillips, W., & Milner, R. M. (2021). *You Are Here. A Field Guide for Navigating Polarized Speech, Conspiracy Theories, and Our Polluted Media Landscape*. Cambridge: The MIT Press.
- Proctor, R. N., & Schiebinger, L. (2008). *Agnotology. The Making and Unmaking of Ignorance*. Stanford: Stanford University Press.
- Quattrociocchi, W., Scala, A., & Sunstein, C.R. (2016). Echo Chambers on Facebook. *Economics of Networks eJournal*. <http://dx.doi.org/10.2139/ssrn.2795110>
- Rini, R. (2017). ‘Fake News and Partisan Epistemology’. *Kennedy Institute of Ethics Journal* 27(2):E-43. <https://doi.org/10.1353/ken.2017.0025>.
- Riemer, K., & Peter, S. (2021). Algorithmic audiencing: Why we need to rethink free speech on social media. *Journal of Information Technology*. <https://doi.org/10.1177/02683962211013358>
- Rosenberger, R. (2019). The Experiential Niche: Or, on the Difference Between Smartphone and Passenger Driver Distraction. *Philosophy & Technology*, 32, 303–320. <https://doi.org/10.1007/s13347-017-0297-8>
- Ruby, T. F. (2006). Listening to the Voices of Hijab. *Women’s Studies International Forum*, 29, 54–66. <https://doi.org/10.1016/j.wsif.2005.10.006>

- Smart, P. (2017). Extended Cognition and the Internet. *Philosophy & Technology*, 30, 357–390. <https://doi.org/10.1007/s13347-016-0250-2>
- Smart, P., Clowes, R., & Heersmink, R. (2017a) ‘Minds Online: The Interface between Web Science, Cognitive Science and the Philosophy of Mind’. *Foundations and Trends® in Web Science* 6(1–2):1–232. <https://doi.org/10.1561/18000000026>.
- Smart, P., Heersmink, R., & Clowes, R. W. (2017b). The Cognitive Ecology of the Internet. In S. J. Cowley and F. Vallée-Tourangeau (Eds.) *Cognition Beyond the Brain: Computation, Interactivity and Human Artifice*, pp. 251–82 Cham: Springer International Publishing.
- Smith, N., & Graham, T. (2019). Mapping the Anti-Vaccination Movement on Facebook. *Information, Communication & Society*, 22(9), 1310–1327. <https://doi.org/10.1080/1369118X.2017.1418406>
- Sousa, R. (2009). Epistemic Feelings. *Mind and Matter*, 7(2), 139–161. Available at: <https://www.mindmatter.de/resources/pdf/desousawww.pdf>
- Sommers, T. (2009). The Two Faces of Revenge: Moral Responsibility and the Culture of Honor. *Biology and Philosophy*, 24, 35–50. <https://doi.org/10.1007/s10539-008-9112-3>
- Srnicek, N. (2017). *Platform Capitalism*. Cambridge; Malden: Polity Press.
- Sullivan, S., & Tuana, N. (2007). *Race and Epistemologies of Ignorance*. SUNY Press.
- Sunstein, C. R. (2001.) *Echo Chambers: Bush V. Gore, Impeachment, and Beyond*. Princeton: Princeton University Press.
- Sunstein, C. R. (2017). *#Republic*. Princeton University Press.
- Sunstein, C. R., & Vermeule, A. (2009). Conspiracy Theories: Causes and Cures\*. *Journal of Political Philosophy*, 17(2), 202–227. <https://doi.org/10.1111/j.1467-9760.2008.00325.x>
- Szabla, M. & Blommaert, J. (2020). Does context really collapse in social media interaction?. *Applied Linguistics Review*, 11(2), 251–279. <https://doi.org/10.1515/applirev-2017-0119>
- Tanesini, A. (2016). I—“Calm Down, Dear”: Intellectual Arrogance, Silencing and Ignorance. *Aristotelian Society Supplementary*, 90(1), 71–92. <https://doi.org/10.1093/arisup/akw011>
- Terpe, S. (2016). Sentiments Epistemològics En Les Experiències Morals i Dinàmiques Morals En La Vida Quotidiana. *Digithum*, 18, 5–12. <https://doi.org/10.7238/d.v0i18.2874>
- Verbeek, P. P. (2005). *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Pennsylvania State University Press
- Verbeek, P.P. (2008) Cyborg intentionality: Rethinking the phenomenology of human–technology relations. *Phenomenology and the Cognitive Sciences* 7:387–395. <https://doi.org/10.1007/s11097-008-9099-x>
- Verbeek, P.P. (2011). *Moralizing Technology: Understanding and Designing the Morality of Things*. University of Chicago Press.
- Vitak, J. (2012). The Impact of Context Collapse and Privacy on Social Network Site Disclosures. *Journal of Broadcasting & Electronic Media*, 56(4), 451–470. <https://doi.org/10.1080/08838151.2012.732140>
- Waite, C., & Bourke, L. (2015). Using the Cyborg to Re-Think Young People’s Uses of Facebook. *Journal of Sociology*, 51(3), 537–552. <https://doi.org/10.1177/1440783313505007>
- Williams, H. T. P., McMurray, J. R., Kurz, T., & Hugo Lambert, F. (2015). Network Analysis Reveals Open Forums and Echo Chambers in Social Media Discussions of Climate Change. *Global Environmental Change*, 32, 126–138. <https://doi.org/10.1016/j.gloenvcha.2015.03.006>
- Wollebæk, D., Karlsen, R., Steen-Johnsen, K., & Enjolras, B. (2019). Anger, Fear, and Echo Chambers: The Emotional Basis for Online Behavior. *Social Media + Society* 5(2):2056305119829859. <https://doi.org/10.1177/2056305119829859>.
- Woods, J. (2005). Epistemic Bubbles. Pp. 731–74 In Artemov, S., H Barringer, A. d’Avila Garcez, L. C. Lamb, and J. Woods. (Eds.) *We Will Show Them! Essays in Honour of Dov Gabbay, Volume Two*, pp. 731–74. London: College Publications.
- Yardi, S. and boyd, d. (2010). Dynamic Debates: An Analysis of Group Polarization over Time on Twitter Bulletin of Science, Technology & Society 30 5 316 327 <https://doi.org/10.1177/0270467610380011>
- Nelson, L. (2018). *Social Media and Morality. Losing Our Self-Control*: Cambridge University Press.

## Authors and Affiliations

**Giacomo Figà Talamanca**<sup>1</sup>  · **Selene Arfini**<sup>2,3</sup>

<sup>1</sup> Department Ethics and Political Philosophy, Radboud University, Nijmegen, The Netherlands

<sup>2</sup> Department of Humanities - Philosophy Section, University of Pavia, Pavia, Italy

<sup>3</sup> Computational Philosophy Laboratory, University of Pavia, Pavia, Italy