Group polarization is a serious and worrying phenomenon developing in democratic societies. It occurs «when members of a deliberating group move toward a more extreme point in whatever direction is indicated by the members' predeliberation tendency» (Sunstein, 1999: 3-4). As a result, after a discussion where different points of view emerge, people tend to align to the opinion they were already tending to, *before* the discussion itself. For example, if confronted with someone that supports a different position, people who believe that vaccines are linked to autism will adhere to their pre-existing perspectives with much more conviction; the same will occur between left-leaning voters or pro-feminist activists and so on.

This mechanism – that creates extreme views in a group after deliberation – clearly leads to a strong fragmentation on political and social issues and, in some cases, to the rise of extremism and fanaticism.

When society develops into factions that are not able to communicate and understand each other, democratic institutions weaken, because democracy needs a healthy public debate (*i.e.* open, constructive, balanced) to remain strong and vital.

Polarization is not a new problem, nor is a specific issue of the digital age. Groups tended to be polarized far before the rise of the Internet and social media.

How is polarization shaped by the peculiar characteristics of Web and digital environments, though? Is the Internet causing a more polarized public opinion, exacerbating differences and confrontations between different groups?

Through the review of the current state of the art, I will try to point out the main characteristics and issues of online polarization. I will refer particularly to the theories developed by Cass Sunstein, one of the first academic that has recognized the importance of group polarization in democratic societies. Then, I will analyse lots of empirical studies that have tried to examine the behaviour of online group users with the aim to give a precise representation of the phenomenon. Moreover, online polarization will be addressed as a result of three different causes interacting with each other:

1. the individual one, represented by bias and heuristics (*i.e.* mental shortcuts based on

empirical thinking);

2. the social one, through the formation of echo-chambers;
3. the technological one, with the rise of algorithms and the so-called *filter bubbles*.

Three Explanations for Polarization

Being polarized doesn't simply mean having strong disagreements with others who share a different worldview: in a free society finding contrasting opinions is a sign of good health, not a manifestation of decadence. Diversity is a desirable virtue for a democratic society. But polarization goes far beyond this natural co-existence of divergent thoughts: it means having personal and «emotionally charged negative feelings» (Blankenhorn, 2015) towards those who think in other ways, that are recognized as members of an opposite, rival group. It's "us" against "them", in which "them" is the enemy, viewed as a group of people who are certainly wrong in their values and beliefs (while "we" are certainly right).

Why does this happen?

There are mainly three reasons why groups tend to polarize:

1. persuasive arguments and information: people should change their minds according to the most convincing argument. But «if the group's members are already inclined in a certain direction, they will offer a disproportionately large number of arguments tending in that same direction, and a disproportionately small number of arguments tending the other way» (Sunstein, 2017: 72). The limited argument pool will only reinforce pre-existing convictions, leading to group polarization;
2. reputational reasons: people care about their reputation and want to be accepted by other group's members. That's why they tend to be aligned to the dominant position, while minority opinions dissolve in a «spiral of silence» (Noell-Neumann, 1984);
3. confidence, extremism and corroboration: polarization increases when people feel more confident. If someone is not certain about an issue, then there's less possibility that he might develop extreme beliefs about that issue, while extremism fosters polarization. Corroboration and agreement from other group's members can increase

someone's confidence: like-minded people talking to each other become more convinced of their opinions and thus more extreme.

It's also important to underline the role of traditional and social media in the development of polarization, as they are vectors of information and places that gather together homogeneous groups of people. News in newspapers, television and digital media can be reliable, truthful, or correct but never completely objective: there will always be a point of view, a shade of interpretation, a trace of subjectivity. Consumers and users choose the source of information that better fits in with their worldviews. Thus, Christians will address to newspapers that are near to the thought of the Church; right-wing voters to the ones that support the traditional values of the Right and so on.

Media, if devoted to one cause and openly partisan, can enhance polarization. This happened also in the past; but how has the Interned changed the situation?

Internet and the "Daily Me"

With the rise of the Internet, and in particular of the so-called Web 2.0 era (O' Reilly, 2004), there has been a real shift of paradigm in the creation and consumptions of information, passing from a mediated (think about the traditional role of journalists) to a more disintermediate selection process. Users are more independent in choosing their source of information, because of the wide availability of contents on the Internet. Furthermore, on social media the exchange of info, from the producer to the consumer is much more direct and interactive: users can debate, react and contribute with their role to the success of a specific content. They can even create new contents, not being just a passive audience but turning into "prosumer" (Toffler, 1980).

In 1995, Nicholas Negroponte theorized the "Daily Me" (Negroponte, 1995), a personalized and tailored made news package, containing only information previously chosen. We are not far from that. On the Internet, users have the power to filter what they see: they can subscribe to specific newsletters or channels (for example on YouTube) and deliberately avoid source of information whose values they don't agree with. These developments

increase individual power and entertainment. Filtering, by itself, is a normal and democratic process: in a free society no one is forced against his or her will to watch or read something. The present situation combines in worrying ways «a dramatic increase in available options, a simultaneous increase in individual control over content, and a corresponding decrease in the power of general interest intermediaries» (Sunstein, 2007: 8). By "general interest intermediaries" Sunstein means newspapers, magazines, broadcasters which may bring to people not only the information they already look for, but also *unwanted and unexpected information.*

How does such condition affect polarization?

First, people on the Web can easily get connected to people with similar views: in like-minded groups opinions tend to polarize. In such environments, it's also easier for individuals to self-segregate ideologically, building gated communities.

As a second issue, users of the social media are predisposed to look for information that confirms their system of beliefs. This could lead to extremism and a distorted vision of reality.

Of Biases, Echo-chambers and Filter Bubbles

I will now try to analyse the three different levels that combined together lead to group polarization on social media.

In spite of the theory of communicative rationality theorized by Apel and Habermas, human beings are not completely driven by rationality when they deliberate and take decisions. Human minds are regulated by two different thinking systems: System 1, dedicated to fast thinking and guided by emotional response and heuristics and System 2, connected to slow thinking, responsible of critical and logical speculation and needing more attention and cognitive effort (Kahneman, 2011). Each of them should be activated by specific situations. Heuristics, which are not linked to logical thinking but use empirical methods to decode reality, are useful *only* in certain circumstances: otherwise they can bring to distortions and

misjudgements. Together with heuristics, individuals are affected in their excogitations by inclinations and prejudices, called bias, determined by personal experience and history, intimate beliefs, personality traits, and the socio-cultural context in which someone lives in.

There are many types of bias: confirmation bias is the most important aspect to be able to understand the role of personal inclinations of polarized communities online. It occurs when «one selectively gathers, or gives undue weight to, evidence that supports one's position while neglecting to gather, or discounting, evidence that would tell against it» (Nickerson, 1998). Other two biases are important for the dynamics of polarized community: the bias blind spot (the inability to detect one's own biases; Pronin, Lin, and Ross, 2002) and the bandwagon effect (individuals adopting the majority opinion; Nadeau, Cloutier, and Guay, 1993).

Personal inclinations can easily lead to the polarizations of one own's beliefs: together with social influence (*i.e.* the process under which someone's values, behaviours or opinions are affected by others), biases can result in polarization (Del Vicario, *et al.*, 2016a; Bessi, Zollo *et al.,* 2015a). This is what I describe as "individual level": opinions tend to polarize because individuals are deceived by cognitive biases. During a discussion, participants will take as reliable only the arguments that confirm their pre-existing views, according to their system of beliefs, and reject those who contradict prior preconceptions. When in Internet, people are exposed to plenty of different opinions, but they will be naturally inclined to listen to the ones with which they already agree. This also proofs that «individuals who receive unwelcome information may not simply resist challenges to their views. Instead, they may come to support their original opinion even more strongly … a "backfire effect"» (Nyhan, Reifler 2010).

This kind of automatisms (bias and heuristics), which have existed since the dawn of times, have intertwined with digital technology in an alarming way, amplifying the diffusion of misinformation and disinformation and poisoning public debates in social networks.

The individual level of polarization on social media merges together with the social one.

Online spaces take shape as groups of social networks and subnetworks (Castells, 1996). These networks are mostly composed by people which share similar values and interests:

this peculiarity is called *homophily*. That's why on Facebook our "friends" are likely to be ideologically and politically homogeneous to us; that's why on Twitter we follow mostly people which we identify with. «Similarity breeds connection» (McPherson, Smith-Lovin and Cook, 2001: 415). The tendency towards homophily is one of the causes of the formation of echo-chambers online (Bright, 2017; Barberá, 2017).

Echo-chambers are homophile clusters, digital environments that gather together homogeneous groups of people. Cass Sunstein has been one of the first academic to study them. On 2001, before the birth of social media, he wrote about the risk of fragmentation brought by these digital spaces, considering it a danger for democracy:

...it is important to realize that a well-functioning democracy—a republic—depends not just on freedom from censorship, but also on a set of common experiences and on unsought, unanticipated, and even unwanted exposures to diverse topics, people, and ideas. A system of "gated communities" is as unhealthy for cyberspace as it is for the real world. (Sunstein, 2001: 2).

Many studies have proved and investigated the existence of echo-chambers on different social networks (about Facebook: Bessi, *et al.,* 2015a; Bakshy, Messing, and Adamic, 2015; Del Vicario, *et al.,* 2016b; Quattrociocchi, Scala, and Sunstein, 2016; Bessi, *et al.* 2016; about Twitter: Himelboim, McCreery, and Smith, 2013; Colleoni, Rozza, and Arvidsson, 2014; Garimella, *et al.,* 2018).

It's not easy in echo-chambers to run into people that have different values and beliefs, because of their homophily. Inside of them, groups tend to polarize because some information and opinions are constantly echoed and repeated while others seem to disappear.

Group's members see only what confirms their previous conceptions and live closed in «information cocoons» (Sunstein, 2017). Moreover, some studies have proved that members of groups that support opposite narrative (such as science vs conspiracy groups) are inclined to interact only with their own community while when they relate to others from different echo-chamber they do not communicate, but express only negative feelings or comments (Bessi, *et al.,* 2015c; Zollo, *et al.,* 2017).

Thus, all these elements put together – confirmation bias, homophily and the isolated and deformed reality of echo-chambers – lead to group polarization. It is also true that some researchers have found that not everyone is using the Internet to segregate in echo-chambers but only those who are already extremist and partisan in their views (Gentzkow, and Shapiro, 2011; Guess, 2016), although these subjects «exercise disproportionate influence» over the political system (Nyhan, 2016).

In summary, the technological element adds heavily to the mixture of biases and echo-chambers. Social media and search engines operate through algorithms, automatic procedures that using a predetermined number of steps aim to solve a problem. These technologies are mainly used for researching, content promotion, selection and filtering. Without algorithms, the average user would be lost in the depths of the Internet.

There are also collateral effects, that must be taken into consideration since they can lead to further distortions and deformations of the real, stimulating polarization and strengthening the effects of bias and echo-chambers (Flaxman, Goel, and Rao, 2016; Bessi, *et al.*, 2016; Johnson, *et al.*, 2017). Eli Pariser, one of the first academics to talk about this phenomenon, created the term filter bubble to address this condition (Pariser, 2011). Algorithms, by giving priority to certain contents, would isolate people in their own bubble, where alternative points of view struggle to enter. This happens because they «foster personalized contents according to user tastes—i.e. they show users viewpoints that they already agree with» (Bessi, *et al.*, 2016). That's exactly what is stated by a post from Facebook Newsroom about changes in the NewsFeed, published on April 21, 2016:

…we make updates to help make sure you see the most relevant stories at the top… With this change, we can better understand which articles might be interesting to you based on how long you and others read them, so you'll be more likely to see stories you're interested in reading (Blanc, Xu, 2016).

Online platforms, through customization and the use of algorithms, give priority to the most relevant contents, but the logic behind this concept is questionable. What does the word relevant mean in such a contest? Is something relevant just because people would be interested in reading it? These tailored made services can bring to polarization because they corroborate people's biases, by giving them what they already like or have searched for.

Moreover, they reinforce the segregation of echo-chambers.

Conclusion

The most serious aspect of the "algorithm dilemma" is that the mediation operated by platforms to the most is unknown: many people do not even imagine according to which criteria Google or Facebook propose contents. In the same way, it's difficult to recognize your own bias or the dynamics of echo chambers. Thus, people believe they are objective and well-informed, while their visions of the world are affected and influenced by mechanisms invisible to them. This makes emancipation even more difficult.

A homogeneous and cohesive group, isolation, and a little presence of contrasting points of view: these are the conditions that lead to group polarization. Polarization is well represented both offline and online. On the Internet, three different dynamics promote the development and spread of group polarization: personal leanings (individual level), the forming of gated communities called echo-chambers (social level) and the selective action of automatic algorithms (technological level). These elements could be considered three different kind of biases – «bias in the brain, bias in society, bias in the machine» (Ciampaglia, and Menczer, 2018) – that combined together isolate communities and strengthen fragmentation.

References

Bakshy E., Messing S., Adamic L. (2015), Exposure to Diverse Information on Facebook, *Science*, 348: 1130-1132, doi: 10.1126/science.aaa1160;

Barberá P. (2017), Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data, *Political Analysis*, 23 (1): 76-91, https://doi.org/10.1093/pan/mpu011;

Blankenhorn D. (December 22, 2015), *Why Polarization Matters*, «The American Interest», https://www.the-american-interest.com/2015/12/22/why-polarization-matters/;

Blank M., Xu J. (2016), More Articles You Want to Spend Time Viewing, Facebook Newsroom, https://newsroom.fb.com/news/2016/04/news-feed-fyi-more-articles-you-want-to-spend-time-viewing/;

Bessi A., Zollo F. *et al.*(2015a), Trend of narratives in the age of misinformation, *PLoS ONE*, https://www.researchgate.net/publication/275279899_Trend_of_Narratives_in_the_Age_of_Misinformation;

Bessi A. *et al.*(2015b), *Viral misinformation: the role of homophily and polarization*, WWW '15 Companion: proceedings of the 24th International Conference in World Wide Web, New York: ACM;

Bessi A. *et al.*(2015c), Science vs Conspiracy: Collective Narratives in the Age of Misinformation, *PLoS ONE*, 10(2), doi:10.1371/journal.pone.0118093;

Bessi A. *et al.*(2016), Users polarization on Facebook and Youtube, *PLoS ONE,*11 (8), https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0159641&type=printable;

Bright J. (2017), Explaining the Emergence of Echo Chambers on Social Media: The Role of Ideology and Extremism, *SSRN Electronic Journal*, DOI: 10.2139/ssrn.2839728; Quattrociocchi W., Scala A., Sunstein C., (2016), *Echo-chambers on Facebook*, https://ssrn.com/abstract=2795110;

Castells M. (1996), *The Rise of the Network Society*, Oxford: Blackwell Publishing;

Ciampaglia G. L., Menczer F. (June 20, 2018), *These are the three types of bias that explain all the fake news, pseudoscience, and other junk in your News Feed*, «NiemanLab»: http://www.niemanlab.org/2018/06/these-are-the-three-types-of-bias-that-explain-all-the-fake-news-pseudoscience-and-other-junk-in-your-news-feed/;

Colleoni E, Rozza A., Arvidsson A. (2014), Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Bug Data, *Journal of Communication*, 64: 317-332;

Del Vicario *et al.*(2016a), Modeling confirmation bias and polarization, *Scientific Reports*,7, doi: 10.1038/srep40391;

Del Vicario *et al.*(2016b), Echo Chambers: Emotional Contagion and Group Polarization on Facebook, *Scientific reports*, 6, doi: 10.1038/srep37825;

Flaxman S., Goel S., Rao J. M. (2016), Filter Bubbles, Echo Chambers, and Online News Consumption, *Public Opinion Quarterly*, 80: 298-320, doi:10.1093/poq/nfw006;

Garimella K. *et al.*(2018), Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship, *International World Wide Web Conference Committee*, https://doi.org/10.1145/3178876.3186139;

Gentzkow M., Shapiro J. M. (2011), Ideological Segregation Online and Offline, *Quarterly Journal of Economics*, 126: 1799-1839, http://www.nber.org/papers/w15916;

Guess M. (2016), Media Choice and Moderation: Evidence from Online Tracking Data, https://www.scribd.com/document/323356298/Media-Choice-and-Moderation-Evidence-from-Online-Tracking-Data;

Himelboim I., McCreery S., Smith M. (2013), Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter, *Journal of Computer-Mediated Communication*, 18: 40-60, https://doi.org/10.1111/jcc4.12001

Johnson, N. F. *et al.*(2017), Population polarization dynamics and next-generation social media algorithms,

https://www.semanticscholar.org/paper/Population-polarization-dynamics-and-social-media-Johnson-Manrique/841ee2edac088939cc129c04fc993e4a2db0b8f5;

Kahneman J. (2011), *Thinking Fast and Slow*, New York: Farrar, Straus and Giroux; McPherson M., Smith-Lovin L., Cook J. M. (2001), Birds of a Feather: Homophily in Social Networks, *Annual Review of Sociology*, 27: 415-444;

Nadeau R., Cloutier E., Guay J. H. (1993), New Evidence About the Existence of a Bandwagon Effect in the Opinion Formation Process, *International Political Science Review,* 14 (2), https://doi.org/10.1177/019251219301400204;

Negroponte N. (1995), *Being digital,* New York: Alfred A. Knopf;

Nickerson R. S. (1998), Confirmation Bias: A Ubiquitous Phenomenon in Many Guises, *Review of General Psychology*, 2 (2): 175-220, http://psy2.ucsd.edu/~mckenzie/nickersonConfirmationBias.pdf;

Nyhan B. (September 7, 2016), *Relatively Few Americans Live in Partisan Media Bubble, but They're Influential*, «The New York Times», https://www.nytimes.com/2016/09/08/upshot/relatively-few-people-are-partisan-news-consumers-but-theyre-influential.html;

Nyhan B., Reifler J. (2010), When corrections fail: the persistence of political misperceptions, *Political Behavior*, 32 (2): 303-330;

Noelle-Neumann E. (1984), *Spiral of silence*, Chicago: Chicago University Press;

O' Reilly T. (2004), What is Web 2.0, https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html;

Pariser E. (2011), *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*, London and New York: Penguin;

Pronin E., Lin D. Y., Ross L. (2002), The Bias Blind Spot: Perceptions of Bias in Self Versus Others, *Personality and Social Psychology Bulletin*, 28 (3), https://doi.org/10.1177/0146167202286008;

Sunstein C. (1999), The Law of Group Polarization, John M. Olin Program in L. & Econ. Working Paper, 91, http://nrs.harvard.edu/urn-3:HUL.InstRepos:13030952;

Sunstein C. (2001), *Echo chambers: Bush v. Gore, impeachment, and beyond*, Princeton: Princeton University Press;

Sunstein C. (2007), *Republic.com 2.0*, Princeton: Princeton University Press;

Sunstein C. (2017), *#republic. Divided Democracy in the Age of Social Media*, Princeton: Princeton University Press;

Toffler A. (1980), *The third wave*, New York: Bantam Books;

Zollo F., *et al.*(2017), Debunking in a world of tribes,*PLoS ONE*, 12(7), https://doi.org/10.1371/journal.pone.0181821.

Share this:

Share