Group Epistemology and Structural Factors in Online Group Polarization

Forthcoming in *Episteme*

Draft – Please cite final version

There have been many discussions recently from philosophers, cognitive scientists, and psychologists about *group polarization*, particularly with regards to political issues and scientific issues that have become markers of social identity, such as anthropogenic climate change and vaccine hesitancy (McCright and Dunlap, 2011; Schmidt et al., 2018). While the term has been used in many different ways, here I will take "group polarization" to refer to a number of related phenomena in which members of groups become more extreme in their beliefs, in a direction determined by the perceived average strength of belief within the group, after discussion (Sunstein, 2002)[1]. Online and social media environments in particular have received a lot of attention in these discussions, both because of people's increasing reliance on such environments for receiving and exchanging information, and because such environments often allow individuals to selectively interact with those who are like-minded. My goal here is to argue that the group epistemologist can facilitate understanding the kinds of factors that drive group polarization in a way that has been overlooked by the existing research. Specifically, I argue that

[1] As Bramson et al. (2017) note, the term "group polarization" has been used to refer to a number of related but different phenomena. For ease of discussion, I will narrow the sense of the term to refer to just the phenomena I describe above and not, for example, the sense in which bipartisan political societies have been referred to as "polarized".

polarization can occur in part because of the ways that members of a group treat the *group itself* (as opposed to an individual member within that group) as a source of information, and in doing so makes their own position, as well as that of the group, more extreme. I refer to this as a *structural factor* in driving polarization, as it is a factor that is produced by the general nature of the relationship between a group and its members.

There are, of course, many existing theories positing explanations of how groups polarize. When considering the behavior of individuals within a group, there are two broad types of factors that are typically taken to drive polarization effects: *social* and *informational*. Social factors are ones that impact an individual's belief formation and updating insofar as one engages in comparison to other members within one's group, such that one becomes more extreme in one's beliefs as to maintain a sense of identity within the group, and to avoid rejection by other group members. Informational factors, on the other hand, may involve a rational (or at least not irrational) response to the quantity and order of information and arguments that a member receives from other members within a group, such that one is persuaded by arguments presented by other group members that supports the overall position of the group, resulting in an increase in the average strength of member belief.

However, I argue that structural factors can contribute to explanations of polarization that occurs in an important subset of groups, namely online and social media groups in which little is known about other members within the group, what I refer to as *anonymous, semi-anonymous and pseudonymous (ASAP)* groups. ASAP groups differ from those in which members interact with one another face-to-face, or in which members know a lot about the identities of other members, in important ways: specifically, there is less *social presence* in ASAP groups, and as a result, members of ASAP groups seek out and interpret information differently, especially with

regards to whom they take to be a trustworthy source. This is not to say that such factors operate independently of social and informational ones; rather, I argue that structural factors can be more prominent drives of polarization in ASAP groups. Furthermore, I argue that structural factors can contribute to explanations of polarization not only in ASAP groups, but also in groups generally.

The paper will proceed as follows. Section 1 provides a summary of some of the most widely discussed theories of how social and informational factors cause polarization in groups, including online groups. Section 2 outlines some important differences between types of groups in terms of the extent to which members know the identity of other members, and argues that for ASAP groups in particular, these differences are important for the ways that group members exchange and acquire information. I argue that with less information about the identities of other group members, individuals will rely much more heavily on other markers of trustworthiness, specifically *collective member endorsement*. Section 3 then argues that the group epistemologist can make an important contribution to the discussion by considering how groups can be considered as providers of information in their own right, and how the relationship between individual member and group can help explain how ASAP groups polarize. Finally, Section 4 concludes by considering how structural factors can help explain polarization in groups in general.

1. Polarization and Identity Transparency

While many discussions of polarization concern differences in views on social, political, and scientific issues at the national level, research on group polarization has shown that as groups can come in many different shapes and sizes, so too does polarization occur in many different kinds

of groups. That being said, we can distinguish between two main types of groups that have been the focus of polarization discussions: groups in which members primarily interact with one another *face-to-face*, and those in which interactions occur *online*. For example, in terms of face-to-face groups, Myers and Lamm (1976) discuss the effects that fraternity members have on one another in terms of their increasingly conservative viewpoints over the course of their education, and Sunstein (2002) considers the ways in which a group of university professors can become more extreme in their views concerning affirmative action after discussion. In terms of polarization in online groups, the online environment that has dominated discussions is Facebook (Guerra et al., 2013; Lee et al., 2014; Bessi et al., 2016; Del Vicario et al., 2016): for instance, in addition to studies investigating the conditions that make Facebook groups prone to polarizing (e.g. Schmidt et al., 2018; Garibay et al., 2019), there have been several potential remedies proposed for polarization on Facebook, which often posit that exposing users to information that they would not normally be exposed to can mitigate polarization effects (e.g Munson et al., 2013; Garimella et al., 2017; although Bail et al., 2018, suggest that this strategy can backfire). Contemporary discussions will often involve both face-to-face and online groups. For example, Broncano-Berrocal and Carter (2020) introduce their discussion of group polarization as occurring in a "high number of socially relevant phenomena such as jury decisions, political debates, financial decision-making, extremism, terrorism, and—of course—interaction with like-minded people on social media" (3-4).

In general, the kinds of factors that have been proposed as causes of polarization are the same in both face-to-face and online groups. However, discussion of online group polarization also involves a distinction between the *internal* interactions between individuals in a group, and the *external* conditions that help give rise to those conditions (Prasetya and Murata, 2020). For

instance, one external explanation that has received a lot of attention is that behind-the-scenes algorithms contribute to group polarization by restricting the kinds of information that individuals within the group will receive (Sunstein, 2017). The above cases, however, tend to focus on the internal conditions, ones which involve discussions between individuals. Here I will not have much to say about the external conditions, but will focus on the internal conditions, instead.

As mentioned above, there are two types of internal factors that are typically taken to explain group polarization. The first are *social factors*, of which there are several varieties. For instance, *social comparison* is a process in which a member adjusts the strength of their beliefs after comparing their views to those of other members within the group (Burnstein and Vinokur, 1977; Myers, 1982): in an early meta-analysis, Isenberg (1986) succinctly describes social comparison as involving a process in which "people are constantly motivated both to perceive and to present themselves in a socially desirable light" and that doing so requires an individual to "be continually processing information about how others present themselves, and adjusting his or her own self-presentation accordingly" (1142). Additional social factors involve the concern for *social identity and categorization*, wherein an individual's identification as being a part of the relevant in-group results in their becoming more extreme as a result of a perceived need to maintain a positive conception of their identity as being a member of that group (Brewer, 1979; Turner, 1985; Hogg et al., 1990; Yardi and Boyd, 2010). As Sunstein (2002) puts in, social factors in general drive group polarization as a result of "individuals mov[ing] their judgments in order to preserve their image to others and their image to themselves" (179).

As a result of social comparison and concern for identity and categorization, members of groups will tend to become more extreme in their relevant beliefs in a direction determined by

the perceived group average. The process of individuals adjusting their beliefs in this way is not typically considered to be a rational one: agents will adjust their beliefs in a direction that is more extreme than the available evidence warrants in order to help preserve their sense of identity and status within the group. While some argue that this process can be rational insofar as it is sometimes rational to preserve one's social identity (Sunstein, 2002) it is nevertheless an *epistemically* irrational response to evidence, in that one's strength of belief does not accurately reflect the available evidence (see Olsson, 2013; Pallavicini et al., 2018).

In addition to social factors, *informational factors* have been posited as potential causes of group polarization. These factors involve the kind of information a member receives within a group, and the way in which one receives it. For example, according to the *persuasive arguments* hypothesis, members of groups become more extreme in their beliefs as a result of being presented with arguments from many different members that all support the same view, thus making it appear as though the balance of reasons heavily supports one view over another (Burnstein and Vinokur, 1977; Sunstein, 2002). Other factors such as *information cascades*, in which consistent initial evidence affects the way in which one interprets potentially contradicting evidence one receives later on, can also impact belief forming and updating in ways that result in member beliefs becoming more extreme (Anderson and Holt, 1997). Unlike social factors, adjusting one's strength of belief in response to informational factors is not necessarily epistemically irrational: if all of the available evidence that one receives is in favor of a particular view, for example, it may very well be rational to increase the strength of one's relevant belief in that view.

So far, we have seen examples of polarization involving groups that differ in terms of size, structure, and goals, as well as in terms of whether their members interact face-to-face or in

online environments. Here I will introduce an additional characteristic that can distinguish different kinds of groups, *identity transparency*. As I will use the term here, identity transparency concerns the extent to which members within a group know the identity of other members, such that a group that admits of a high degree of identity transparency will be those in which members know a lot about the identities of other members within the group, while groups that have a low degree of identity transparency will be those in which members do not know much or anything about the identities of other members in the group. There is, of course, a lot of potential variance of identity transparency between groups: for example, while neighborhood communities and juries may tend to be the kinds of groups that have a high degree of identity transparency, I may still know a lot more about my neighbours than I do fellow members of a jury. Some groups may also have a higher potential for identity transparency than others. For instance, while I may have never introduced myself to my colleagues across the hall, the group of colleagues that we are both members of has a high potential for becoming identity transparent, since I am in a position to easily come to know more about them. On the other hand, a group of anonymous strangers that I interact with online will both have a low degree of identity transparency, as well as a lower potential for identity transparency, given that there is no easy way for any of us to learn about the identities of one another.

As we will see in what follows, the degree of identity transparency in both face-to-face and online groups can vary significantly. However, as I show in the next section, identity transparency tends to be lower in online groups: while there is evidence that some types of online conversations, especially political ones, occur most frequently between friends and relatives (Tucker et al., 2018), this is certainly not always the case, as the internet offers users the opportunity to converse with people of all degrees of familiarity. Specifically, I discuss next an

important subclass of online groups in which there is polarization that has a very low degree of identity transparency.

2. Online, Social Media, and ASAP Groups

Call a group an *anonymous, semi-anonymous, or pseudonymous (ASAP)* group just in case it is a group in which the identities of members are either not known to one another (e.g. in which all members of the group are anonymous), are known only in some limited capacity, or are only known to each other only on the basis of pseudonym. As a result of the potential for anonymity, such groups will tend to have very low degrees of identity transparency. While it is not necessary that such groups occur in online environments, I take it that they are most commonly found online: for example, message board websites (e.g. Reddit), and social media in which a user's identity is not required in order to participate in group discussion (e.g. Twitter) are environments in which one can find ASAP groups. In contrast, the kinds of face-to-face groups provided as examples above (e.g. juries, groups of community members, colleagues, etc.) will generally fall outside of this category. This is not to say that all online groups are ASAP groups: as mentioned above, some online groups can have high degrees of identity transparency (e.g. when one discusses political issues with family and friends), and there may very well be face-to-face groups in which members know little to nothing about one another (e.g. the group of criminals in Quentin Tarantino's *Reservoir Dogs* who identify themselves via color-coded pseudonyms). Nevertheless, here I will focus on ASAP groups that occur in online environments. My argument here is that the differences between ASAP groups and face-to-face groups (along with other groups that have high degrees of identity transparency) are important when considering how groups polarize. To illustrate, I will begin by surveying some of the research on the differences

between face-to-face (FtF) communication and computer-mediated communication (CMC) (Siegel et al., 1986; Valacich et al., 1994; Walther and Parks, 2002; Walther, 2011).

Some of the major questions in CMC research concern how personal interactions differ between computer-mediated environments and face-to-face environments, and especially how the effects of online anonymity impact adherence to social norms (Abrams et al., 1990; Marino et al., 2016). One important difference pertains to the ways in which people seek out and interpret information from online groups as compared to face-to-face groups. For instance, Kane et al. (2014) argue that there is reason to believe that the ways individuals interact with others online "may bear little connection to offline social relationships" (286), in that face-to-face group discussions tend to be more goal-oriented, and involve more interpersonal deliberation. Kane and Fichman (2009) also argue that while collaborative online projects like Wikipedia involve the contribution of the work of many individual members, it does not involve the kind of discussion and deliberation between members that one will tend to find in face-to-face groups. Instead, they argue that the ways that individuals seek out information from face-to-face groups tend to involve interactions and discussions with other individuals, whereas when seeking information online one will tend to rely less on interpersonal deliberation and more on *aggregate member endorsement* (Kane and Fichman 2009). This kind of behavior is especially prevalent when making decisions about purchasing products: if I were to seek out information from people in face-to-face groups, I would likely consult a number of different people, and discuss their choices with them. In online groups, however, I am much more likely to simply look for the product with the highest aggregate rating (ibid.).

A key reason for the difference for the way that individuals acquire information when face-to-face as opposed to online is that there is less *social presence* in online environments,

where social presence is defined as "the degree to which people establish warm and personal connections with each other in a communication setting" (Sia et al., 2002: 73; see also Short et al., 1976). As Sia et al. (2002) argue, social presence is important in determining the ways in which individuals seek out and interpret information from others, and is characterized by three types of communication cues: verbal cues that pertain to tone, volume, and rate of speech (see also Cook and Lallijee, 1972; Daft et al., 1987; McGrath, 1984); visual cues, including facial expressions and body language; and textual cues, which pertain to the information included in written text. Importantly, Sia et al. argue that computer-mediated communication generally involves lower social presence than face-to-face communication (see also Poole and Jackson, 1993), and that anonymity lowers it further. One reason is that "communication cues that typically yield higher social presence are those that convey immediacy" (where "immediacy" is defined as "the psychological distance between people who are communicating"), and the kinds of communication cues that one receives in face-to-face communication – e.g. verbal and visual cues – are ones that convey immediacy, whereas the cues available in anonymous computer-mediated communication – e.g. textual cues – do not (Sia et al., 2002: 74).

Lower social presence in online communication has important consequences for thinking about polarization. The first is that in online settings in which one can be anonymous, individuals are more open about sharing their views, presumably because they do not feel apprehensive about being criticized for them (Nunamaker et al., 1991). Interestingly, Sia et al. (2002) argue that a lack of apprehension will result in the production of more novel arguments in ASAP groups, which will in turn result in greater polarization effects, as "people tend to focus on arguments rather than presenters", and will thus try to outdo other members in a process of "one-

upmanship", where members defend views that are more extreme in accordance with the perceived values of the group (78).

Secondly, and more importantly for my purposes here, is that with lower social presence and fewer communication cues, individuals in online groups will look to different types of markers of trustworthiness than in face-to-face groups. As mentioned above, one of the most prominent additional cues comes in the form of *endorsement*: this may come in the form of explicit endorsement markers – e.g. "likes", "hearts", or "upvotes" (Willemsen et al., 2012) – or more indirect cues, such as the number of connections that one has within a given network – e.g. the number of "friends" one has on a social media network (Lim and Van Der Heide, 2015). In general, information that has been highly endorsed is more readily accepted, and members who are highly endorsed by other members tend to be seen as more trustworthy (Willemsen et al. 2012)[2].

The reliance on endorsement as a marker of trustworthiness is the result of one having only limited social cues upon which to evaluate others. Metzger et al. (2010), for instance, found that "participants developed strategies to assess a source's credibility as best they could" (421), given both the overall dearth of information about other members, and the concern that what information might be available – say, in the form of information presented on user-generated

[2] This is not to say that endorsement correlates perfectly with perceived trustworthiness. Willemsen et al. (2012), for example, argue that individuals evaluating the trustworthiness of others online face the "authenticity dilemma": while being very highly endorsed is taken to be a sign of trustworthiness, *universal* endorsement is often taken to be a sign of untrustworthiness and manipulation.

profiles – could be manipulated and curated by the members themselves. In order to evaluate the credibility of other members, then, individuals will tend to employ what Walther et al. (2009) call "warranting theory": in determining the credibility of others, individuals will seek out credentials that are the least susceptible to manipulation. In online environments, such credentials are most readily available in the form of aggregate ratings from the group: this is because as information aggregates it becomes more and more difficult to manipulate, and any potential subjective biases in the ratings of individuals and information will have less of an effect on endorsement overall (Flanagin and Metzger, 2013). Aggregate endorsement can also be taken to be a marker of expertise, with individuals endorsed by members being deemed more credible than those who are self-proclaimed experts (Willemsen et al., 2012).

What research on communication and social cues in online communication suggests, then, is that without the kinds of social cues one can rely on to help assess credibility in face-to-face interactions, one's interactions with others in online groups will tend to be mediated in different ways, with there being an especially significant role for endorsement. The extent to which some information or individual is endorsed, however, is something that is not determined at the level of the individual, but is instead determined at the level of the *group*: that information one provides has received a significant amount of "likes" or "hearts" is not information that is provided by an individual member, but is instead the result of an aggregate of actions from multiple members. That this is the case implies that there is an important role that the information one receives from the *group itself* plays in the way that individual members form and update beliefs in online environments. In the next section I argue that this relationship between individual member and group implies that there are additional *structural factors* that can cause polarization in ASAP groups.

3. Structural Polarization in ASAP Groups

In this section I argue that group polarization in ASAP groups can come about at least in part because of the relationship between members of the group and the group itself, specifically when members appeal to groups as sources of information. For this argument to be successful I will defend three views: first, that online groups themselves (in addition to individual members within a group) can be sources of information; second, that one way that online and ASAP groups can provide information is via the collective endorsement of its members; and third, that relying on and contributing to that endorsement makes it more likely that members within the group will become more extreme in their beliefs. Defending these views will require looking in more detail at some recent work concerning the epistemology of groups.

3.1 Online Groups as Sources of Information

The first view that needs defending is that online groups, in addition to individual members of such groups, can be sources of information. Questions concerning how groups can possess beliefs and knowledge, as well as how they can make assertions and be sources of testimony, have been matters of considerable discussion in recent work in social epistemology. One of the main motivations for positing that groups can be sources of information pertains to common practices of information-seeking. For example, we seek out the national weather service to form beliefs about what the weather will be like tomorrow; we seek out NASA to get the latest news on exoplanet discoveries; we seek out laboratories to learn about new developments in drug research; etc. (Tollefsen, 2007; Fricker, 2012; Lackey, 2018). The manner in which groups provide information will depend on its structure: for instance, groups might be structured in such a way that an individual spokesperson delivers information on the group's behalf (Lackey,

2018), or perhaps members of a group might all contribute to a study or otherwise collaborate on a larger project (Fricker, 2012), etc.

What information a group can provide is a matter of debate in the epistemology of groups. In general, the information that a group can provide is determined by its members; that being said, there are two competing theories about how the states and actions of a group's members determine those of the group itself. According to the *summativist*[3], a group provides information that *p* just in case most or all of its members provide or are in a position to provide that information (Lackey, 2014). For example, a group might take a vote among its members as to its position on *p* and release a statement to that effect, in which case the group provides information that *p* in virtue of the fact that most or all of its members provide that information as well. According to the *non-summativist*, a group can provide information that few or none of its members provide or are in a position to provide (Tollefsen, 2007). For example, in a group of individuals collaborating on a large project it can be the case that while all of the individuals contribute towards figuring out that *p*, only very few of them are actually in a position to provide the information that *p*. I will not here take a stance on the theoretical debate between the summativist and non-summativist views. What is important for my purposes is that however we think about groups as being able to provide information, the information that a group provides is determined, in some way, by its members.

[3] Note that the terms "summativism" and "non-summativism" are used to identify a number of different positions within group epistemology, both in terms of group doxastic and epistemic states, as well as group actions. The senses in which I am using the terms here comes from discussions of group assertion and testimony (Tollefsen, 2007; Lackey, 2014).

It is becoming increasingly common for individuals to seek information from social media groups. For instance, a 2018 PEW Research Center study found that 20% of Americans reported that they often use social media as a source of news, surpassing those who often get their news from newspapers (PEW, 2018). Individuals seem to be able to acquire information from online and social media groups in different ways. For instance, one might treat an online group as an environment in which to seek out information from other individuals, e.g. I might use Twitter to learn about some recent event in the news, but in doing so I use Twitter as a platform through which I can communicate with other individuals. In these cases, I am still relying on individuals for information, they just happen to be individuals that are members of the same group as me. It also seems that I can, however, treat online and social media groups as sources of information in their own right. This is perhaps most clear in the case of sites like Wikipedia, where information is provided as the result of collaboration amongst many individuals: when one treats a Wikipedia entry as a source of information, then, one is relying not on any individual agent, but a group (Tollefsen 2009).

While there is much to be said about these and related debates in the epistemology of groups, what we can conclude from this section is that there is a good theoretical basis for thinking that groups can be sources of information or, at the very least, that groups are treated as sources of information. With that being said, I turn to the next view in need of defense, namely that one way in which groups can provide information is via aggregate member endorsement.

3.2. Aggregate Endorsement as Group Information

Consider the following example of a case in which one might acquire information from an online group:

Healthy Food News: Robin is a member of the "Healthy Food News" Facebook group, and visits it often to get the latest news about healthy food. One day, she sees a highly endorsed post declaring that "açai berries are a new superfood", and that they are so good for you because "they are full of antioxidants". Robin then comes to believe that açai berries are a new superfood, and adds her endorsement to the post. Later, when someone questions whether açai berries are really all they are hyped up to be, Robin makes a post responding that they are great because of how many antioxidants they have. Other members within the group respond similarly, resulting in a group consisting of members who are firm believers in the benefits of açai berries.

We have seen that while groups may provide information via the collaboration of individuals, or just as the product of information provided by the majority of its members, the above example illustrates what I argue is a common additional way in which online groups can provide information, namely via the collective endorsement of its members[4]. We should think that highly endorsed information is provided by the group (in addition to the individual member) for two reasons: first, that collective endorsement is akin to a kind of voting, where a group's position is determined by the collective actions of its members; and second, what information one is able to receive within the group is a product of how highly endorsed it is, a consequence being that the group itself is a source of that information. I defend these views in turn.

[4] I do not take these to exhaust the ways in which one can acquire information from online groups. For instance, such groups can provide information in the form of group mandates, rules, announcements, and FAQs, etc. Thanks to an anonymous reviewer for pointing this out.

First, we have seen above that endorsement is taken by online users to be a particularly salient mark of trustworthiness, given the general lack of relevant available communication cues in online environments. However, given that collective member endorsement is a phenomenon that occurs at the level of the group, information that is accepted because it is highly endorsed is not necessarily accepted solely on the basis of any characteristics of the individual who initially provides that information, but rather because of a characteristic of the group. In this way, we can conceive of the group itself as being a provider of the relevant information[5]. Consider again one of the ways that groups can be sources of information mentioned above, namely in terms of taking a vote of the membership. While a certain motion may be proposed by an individual member of that group, that the majority of members of that group approve of it results in it being the position of the group itself. Collective endorsement can then be seen as a kind of voting: that a significant number of members have shown their approval of some information results in that information being the position of the group, as well. When that information is displayed along with the fact that it has been so highly endorsed, then, is a way in which a group can be said to also be providing that information.

[5] Note that in accepting some highly endorsed information one may very well also take the characteristics of the initial provider of that information into account, especially if relevant facts about that person's identity is known (e.g. if they are a known expert on a given matter). However, when this information is absent (as will often be the case in ASAP groups), characteristics of the initial provider of information will play less of a role in one's accepting that information. Thanks to an anonymous reviewer for suggesting I clarify this point.

The second reason to think that highly endorsed information is information one acquires from the group, is that endorsement determines which information one will be exposed to in said group. That the information one receives is determined by group endorsement is a standard feature of social media sites like Facebook and Twitter: for example, Facebook's algorithms determine which stories appear on one's timeline at least in part due to "the number of comments, likes and reactions a post receives and what kind of story it is"[6], and Twitter's algorithms select "each Tweet using a variety of signals, including how popular it is and how people in your network are interacting with it"[7]. Again, the fact that what information is disseminated to members is determined by the group-level phenomenon of collective member endorsement is then another reason to think that said information is provided by the group itself.

To say that when information is highly endorsed that it is thereby information presented by the group is not to remove the need for the input of any individual member. For example, in the *Healthy Food News* case, there was an individual member who was responsible for making the initial post asserting that açai berries were a superfood. However, given the lack of communication cues and low degree of identity transparency, Robin would have little to go on if she were to receive that information from the individual. That the information is highly endorsed, however, not only made it possible for Robin to receive it in the first place, but provides trustworthiness cues that provides her with a good basis for accepting it. Given that she knows little or nothing about the individual providing the information implies that she is not relying on that individual for said information, but is instead relying on the group itself.

[6] https://www.facebook.com/help/166738576721085

[7] https://help.twitter.com/en/using-twitter/twitter-timeline

What this discussion shows is how work from group epistemology can help us make sense of how online groups can provide information as a product of collective member endorsement. As we saw above, collective action on the part of individuals within a group can result in the group providing information that would not be able to be provided by any individual member within the group. Given the role that endorsement plays in ASAP group as a marker of trustworthiness in the absence of other cues, then, it seems that in such groups, members will often rely on collective member endorsement as a way to acquire information from the group. Next, to show how such views can result in group polarization, we need to look at the relationship between members and groups to see how receiving information from groups via collective member endorsement, while also contributing to it, can result in polarization effects.

3.3 Collective Member Endorsement and Group Polarization

Reliance on groups as a source of information in the form of aggregate endorsement demonstrates an important kind of relationship that occurs within groups, namely one that occurs between a member and the group itself. To see how this relationship can result in group polarization, we need to recognize how members both acquire information from and contribute to the information that is provided by a group. Consider again the *Healthy Food News* case: the view that açai berries are a superfood because they are high in antioxidants is highly endorsed by the members of the group, and thus when Robin checks the group, she will both be more likely to be exposed to that information because it has been so highly endorsed (as we saw above, social media tend to organize information such that highly endorsed information is more prominently displayed) and because in online environments with fewer trustworthiness cues, will be more likely to accept that information on the basis of that endorsement. If Robin acquires a new belief that açai berries are a superfood, then she will contribute to the potential polarization of the

group, in that the average strength of belief in the relevant matter will increase. If this is a belief that Robin already held previously, then it seems likely that she would increase the strength of her belief on the basis of that endorsement, as one learning that one's beliefs are widely approved of is likely to make one more confident that one is correct (Del Vicario et al., 2016; Bessi et al., 2016). Again, an important part of how Robin acquires or updates her belief is due to the relationship between her and the group itself: that she receives the information at all is the result of the collective endorsement of the members, and part of the reason why she accepts the information is because it has been so highly endorsed.

At the same time, we have seen that what information a group can provide is a function of the views and actions of its members. As one receives information from a group via collective member endorsement, then, so too can a member contribute to that endorsement. Indeed, this seems to often be what occurs when considering endorsement behavior in social media groups. For instance, Brandtzaeg and Haugstveit (2014) found that one of the main motivations behind Facebook users "liking" posts in groups with humanitarian causes was that the content was something that they wanted to share with others, while Huang (2013) found similar results when users were interacting with a brand's Facebook group, and Guy et al. (2016) found that across multiple different kinds of social media platforms that the motivation for liking behavior was dominated by factors such as "I learn something from it" and "I agree with it". If these are the kinds of motivations behind endorsement behavior, then it seems that when acquiring information from highly endorsed posts one will likely contribute one's endorsement to it, so long as it is something that the user agrees with or wants to disseminate to others.

By relying on and contributing to group endorsement, then, a member of a group can both increase polarization by acquiring or strengthening her own belief, as well as by making it

more likely that other members will do this same: after all, by contributing to group endorsement other members will be more likely to be presented with that information, and with it being even more highly endorsed such members will interpret it as being more trustworthy. When this process occurs with multiple members of the group – say, when more members look at and subsequently endorse the same post – polarization effects can occur: as a group endorses a view more and more strongly, more and more members will in turn either believe that information or believe it more strongly. As a result, the average strength of belief of members within a group will tend to become more extreme, solely by relying on a group as a source of information.

Furthermore, seeking out information from groups in the way described here can result in additional behaviors that can reinforce polarizing effects. For example, while Robin's contributing her own endorsement in the *Healthy Food News* case is by itself enough to contribute to the polarization of the group, her finding the relevant information trustworthy results in her expressing her belief and defending it in response to challenges, which can in turn make it seem more trustworthy, resulting in other members increasing the strengths of their respective beliefs. While these actions need not always occur, they can be the result of a process that started with a member merely relying on, and subsequently contributing to, information acquired from a group. Polarization can occur, then, as a result of the basic fact that groups are structured such that group views are constituted by member views and actions, and that members seek out information from those groups.

Here, then, is a summary of the main argument: in addition to social and informational factors, structural factors contribute to group polarization in ASAP groups. This is because in such groups the low degree of identity transparency and lack of communication cues forces individuals to rely on alternative markers of trustworthiness, the most prominent being collective

member endorsement. However, since collective member endorsement is a phenomenon that occurs at the level of the group and not the individual, the group itself ought to be considered as the source of the relevant information. Furthermore, because of the ways in which members interpret and contribute to group endorsement, highly endorsed information is likely to spread and to increase the confidence in members' beliefs, a process that can result in an overall increase in the average strength of belief in the relevant information. Therefore, the relationship between member and group, and not just the relationships between individual members, is an important one to consider when explaining group polarization.

I have considered this argument as applying to ASAP groups because such groups are ones in which reliance on groups as sources of information is arguably more prevalent than in other types of groups. That being said, to conclude I will consider how we might apply the argument I have developed here to groups in general. In doing so, I will consider how structural factors relate to the social and informational factors discussed above.

4. Structural Polarization in General

With lower social presence and fewer trustworthiness cues in ASAP groups, turning to groups as providers of information via collective member endorsement makes sense as a way to help individuals determine whether they should accept a given piece of information. There is reason to think, however, that members will appeal to groups as a source of information even in face-to-face groups, and even groups with high degrees of identity transparency. For example, when trying to decide on a verdict, members of a jury will periodically take anonymous votes to get a sense of whether the group is leaning towards a verdict of guilty or innocent. Information

received from the group in this way can influence the degree to which a jury member holds their relevant belief, and in turn can affect the view of the group. Thus, it is not only interactions between jurors, but also interactions between individual jurors and the jury as a group that can drive polarization.

This is not to say that structural factors are the sole, or even most prominent drivers of polarization in groups, generally speaking. However, since structural factors appeal to basic relationships between groups and the members that comprise them, there is reason to think that they will have some role to play in polarization of any type of group, ASAP or otherwise. One consequence of the view that structural factors play a role in group polarization is that it may be a basic feature of groups that are treated as a source of information by their members that they will tend to polarize, regardless of the ways in which individual members interact with one another. In contrast, existing studies that model polarization in theoretical groups are ones in which polarization occurs as the result of the exchange of information between individual members, and are often modelled in such a way that every member communicates with every other member (see Olsson, 2013; Pallavicini et al., 2018; Singer et al., 2019). If what I have argued for here is correct, though, then there are factors that can cause polarization that do not require there to be any discussion *between members*: instead, it can occur solely as the result of members both contributing to the position of the group and relying on the group itself as a source of information. Indeed, such factors are important to consider when modelling the kinds of groups that have been my main focus here: while models of Bayesian agents are obviously idealizations, the nature of such models as consisting of individuals constantly sharing information with one another fails to capture the nature of ASAP and social media groups, where information is not accepted directly from other individual group members, but only insofar as it

receives the endorsement of the group. Much of the discussion in such groups, then, may be better described as existing between member and group, and not between individual members.

This is not to say that structural factors necessarily operate independently of social and informational factors. Indeed, as mentioned above, one way in which social and structural factors may work together is in juries, wherein individual jurors not only share information with each other, but periodically check in with the group, as well, primarily as a way of expediting the process of reaching an agreement: in such a case an individual juror may adjust their belief both on the basis of feedback from individuals and the group. Similarly, Myers (1978) found that exposing experimental participants to the average of group judgments had polarizing effects: while Myers postulated that social factors were driving polarization, in that participants wanted to make sure that they adjusted their views in order to be perceived favorably by others, we can see that one of the means to making such a social comparison may be via receiving information from the group. This may especially be the case when the group of which one is a member is particularly important to one's political or social identity: for example, if I am a member of a group with a particular political orientation, the fact that the aggregate of member judgments is strongly in favor of a certain stance on a given issue may cause me to strengthen my belief accordingly, in order to be perceived favorably by the group.

Structural factors, then, may very well operate alongside or reinforce social factors. However, in introducing the concept of ASAP groups, I have attempted to identify a class of cases in which social factors are likely to play a much less significant role in polarization, given that such groups have a low degree of identity transparency and are not necessarily focused on any topic that is a maker of social identity. In these cases, then, I have argued that information from the group in the form of aggregate member endorsement is employed as a trustworthiness

cue, and not necessarily a means of social comparison. Again, there will certainly be cases in which such information can play both roles, as is demonstrated in jury cases and from experimental results.

It should also be noted that appealing to a group for information is not necessarily a bad thing, and that the potentially beneficial effects of exposing individuals to average group responses have also been recognized in other areas of study, specifically in terms of optimal decision-making. For instance, the Delphi method of opinion aggregation (Linstone and Turoff, 1975) involves panels of experts anonymously answering questions, providing feedback to the responses of others, and being given information about the median of responses from other participants, with the idea that the anonymity afforded to group members will reduce any perceived social pressures, and that information about median group response can help individual members and the group as a whole reach better judgments. Of course, a key difference between structured decision-making cases and the kinds of ASAP and online groups that I describe here is that the latter occur organically and in ways that are not necessarily goal oriented, and that participants in the former are well-aware of the aims of explicitly appealing to information provided by the group. Thus, while it has been recognized that average or aggregate group judgment can be a tool to expedite and facilitate decision making, what discussions of structural factors can highlight is how this information can play a different role in ASAP and online groups, and how such factors can drive polarization.

Structural factors also need not operate independent of informational factors: as groups can be sources of information it stands to reason that they can play a role in the effects of informational factors on polarization. For example, information provided by the group may contribute to information cascades: aggregate member views may be considered an additional

piece of information that will lead one to downplay the effect of later, conflicting information; furthermore, one may be presented with group endorsement that becomes increasingly large with each additional contribution of consistent information, thus providing a way in which information cascades could occur primarily by relying on group information[8]. Of course, information cascades do not have to involve any explicit appeal to the group for information, and do not require appealing to any specific average or aggregate of member views. Furthermore, information acquired from groups need not be presented to one in a sequential manner: one can simply receive highly endorsed information from the group immediately without having acquired any other information beforehand. While informational factors like information cascades then appeal to information received from individuals first and foremost, structural factors do not. In general, then, while groups themselves can play a role in both social and informational drivers of polarization, given that there are classes of groups that polarize where such factors are much less significant gives us reason to think that structural factors are not a mere subset of either.

While much has been said about group polarization, my goal here has been to show that such discussions tend to focus on the group as merely a way to cordon off a collection of individuals, and not as something that is capable of providing information in its own right. By looking at work from the epistemology of groups, however, we can see how members of groups do, in fact, rely on groups for information, and how the relationship between group and member can contribute to group polarization. I have argued that while these effects are perhaps most prominent in ASAP groups in particular, and groups with low degrees of identity transparency

[8] Thanks to an anonymous reviewer for this interpretation of information cascades, and for suggesting the comparison in the first place.

more generally, that we should also consider structural factors as potential contributors to group polarization regardless of the type of group involved[9].

Works Cited

Abrams D, Wetherell M, Cochrane S, Hogg M, and Turner J. (1990). 'Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization.' *British Journal of Social Psychology* 29, 97-119.

Anderson L.R. and Holt C.A. (1997). 'Information cascades in the laboratory.' *The American Economic Review* 87(5), 847-862.

Bail C. A., Argyle L.P., Brown T.W., Bumpus J.P., Chen H, Hunzaker M. F., ... and Volfovsky A. (2018). 'Exposure to opposing views on social media can increase political polarization.' *Proceedings of the National Academy of Sciences*, 115(37), 9216-9221.

Bessi A, Petroni F, Del Vicario M, Zollo F, Anagnostopoulos A, Scala A, ... and Quattrociocchi W. (2016). 'Homophily and polarization in the age of misinformation.' *The European Physical Journal Special Topics* 225(10), 2047-2059.

[9] [acknowledgements]

Bessi A, Zollo F, Del Vicario M, Puliga M, Scala A, Caldarelli G, et al. (2016). 'Users Polarization on Facebook and Youtube.' *PLoS ONE* 11(8): e0159641. https://doi.org/10.1371/journal.pone.0159641

Brewer M. B. (1979). 'In-group bias in the minimal intergroup situation: A cognitive motivational analysis.' *Psychological Bulletin*, 86, 307-324

Bramson A, Grim P, Singer D.J., Berger W.J., Sack G, Fisher S, Flocken C, and Holman B. (2017). 'Understanding polarization: meanings, measures, and model evaluation.' *Philosophy of Science* 84, 115-159.

Broncano-Berrocal F and Carter J.A. (2020). *The Philosophy of Group Polarization*. Routledge.

Burnstein E and Vinokur A. (1977). 'Persuasive argumentation and social comparison as determinants of attitude polarization.' *Journal of experimental social psychology* 13(4), 315-332.

Christopherson K. (2007). 'The positive and negative implications of anonymity in Internet social interactions: 'On the Internet, Nobody Knows You're a Dog''. *Computers in Human Behavior* 23, 3038-3056.

Cook M and Lallijee M. (1972). 'Verbal substitutes for visual signals in interaction.' *Semiotica* 6, 212–221.

Daft R. L., Lengel R. H. and Trevino L. K. (1987). 'Message equivocality, media selection, and manager performance: Implications for information systems.' *MIS Quarterly* 11(3), 355–366.

Del Vicario M, Vivaldo G, Bessi A. et al. (2016). 'Echo Chambers: Emotional Contagion and Group Polarization on Facebook.' *Scientific Reports* 6: doi:10.1038/srep37825.

Flanagin A.J. and Metzger M.J. (2013). 'Trusting expert- versus user-generated ratings online: The role of information volume, valence, and consumer characteristics.' *Computers in Human Behavior* 29, 1626-1634.

Fricker M. (2012). 'Group Testimony? The Making of A Collective Good Informant.' *Philosophy and Phenomenological Research* 84(2), 249-276.

Garimella K, De Francisci Morales G, Gionis A and Mathioudakis M. (2017). 'Reducing controversy by connecting opposing views.' In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 81–90.

Guerra P.C., Meira W, Cardie C. and Kleinberg R. (2013). 'A measure of polarization on social media networks based on community boundaries.' In *Seventh International AAAI Conference on Weblogs and Social Media*.

Hogg M.A., Turner J.C. and Davidson B. (1990). 'Polarized norms and social frames of reference: A test of the self-categorization theory of group polarization.' *Basic and Applied Social Psychology* 11(1), 77-100.

Isenberg D.J. (1986). 'Group Polarization: A Critical Review and Meta-Analysis.' *Journal of Personality and Social Psychology* 50(6), 1141-51.

Kane G.C., Alavi M, Labianca G, and Borgatti S.P. (2014). 'What's Different about Social Media Networks? A Framework and Research Agenda.' *MIS Quarterly* 38(1), 274-304.

Kane G.C., Fichman R.G., Gallaugher J, and Glaser J. (2009). 'Community Relations 2.0: With the Rise of Real-Time Social Media, the Rules About Community Outreach Have Changed.' *Harvard Business Review* 87(11), 45-50.

Lackey J. (2018). 'Group Assertion.' *Erkenntnis* 83(1), 21-42.

---. (2014). 'A deflationary account of group testimony.' In J. Lackey (Ed.). *Essays in Collective Epistemology*, pp. 64-94. Oxford: Oxford University Press.

Lee J.K., Choi J, Kim C, and Kim Y. (2014). 'Social Media, Network Heterogeneity, and Opinion Polarization.' *Journal of Communication* 64(4), 702-22.

Lim Y and Van Der Heide B. (2015). 'Evaluating the Wisdom of Strangers: The Perceived Credibility of Online Consumer Reviews on Yelp.' *Journal of Computer-Mediated Communication* 20, 60-82.

Linstone H. and Turoff M. (1975). *The Delphi Method: Techniques and Applications*. Addison-Wesley Publishing Company.

List C and Pettit P. (2011). *Group agency*. Oxford: Oxford University Press.

Marino C, Vieno A, Pastore M, Albery I.P., Frings D and Spada M.M. (2016). 'Modeling the contribution of personality, social identity and social norms to problematic Facebook use in adolescents.' *Addictive behaviors* 63, 51-56.

McCright A.M. and Dunlap R.E. (2011). 'The politicization of climate change and polarization in the American public's views of global warming, 2001–2010.' *The Sociological Quarterly* 52(2), 155-194.

McGrath J. E. (1984). *Groups: Interaction and Performance*. Prentice Hall: Englewood Cliffs, NJ.

Messing S. and Westwood S.J. (2012). 'Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation When Selecting News Online.' *Communication Research* 41(8), 1042-63.

Metzger M.J., Flanagin A.J., and Medders R.B. (2010). 'Social and Heuristic Approaches to Credibility Evaluation Online.' *Journal of Communication* 60, 413-39.

Meyers R. A. and Seibold D. R. (1987). 'Interactional and non-interactional perspectives on interpersonal argument: Implications for the study of group decision-making.' In van Eemeren F.H., Grootendorst R, Blair J.A. and Willard C.A. (Eds.). *Argumentation: Perspectives and Approaches*, 205-214. Foris: Dordrecht, Holland.

Munson S. A., Lee S.Y. and Resnick P. (2013). 'Encouraging reading of diverse political viewpoints with a browser widget.' In *Seventh International AAAI Conference on Weblogs and Social Media*.

Myers D. (1978). 'Polarizing Effects of Social Comparison.' *Journal of Experimental Social Psychology* 14, 554-563.

Myers D and Lam H. (1976). 'The Group Polarization Phenomenon.' *Psychological Bulletin* 83(4), 602-27.

Nunamaker J. F, Dennis A. R., George J. F., Valacich J. S. and Vogel D.R. (1991). 'Electronic meeting systems to support group work.' *Communications of the ACM* 34(7), 40–61.

Olsson E.J. (2013). 'A Bayesian simulation model of group deliberation and polarization.' In *Bayesian argumentation*, 113-133. Springer: Dordrecht.

Pallavicini J, Hallsson B and Kappel K. (2018). 'Polarization in groups of Bayesian agents.' *Synthese*, 1-55.

PEW Research Center. (2018). 'Social media outpaces print newspapers in the U.S. as a news source.' https://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/. Accessed on January 8, 2020.

Poole M. S. and Jackson M. H. (1993). 'Communication theory and group support systems.' In Jessup L. M. and Valacich J. S. (eds.). *Group Support Systems: New Perspectives*, 281–293. Macmillan: New York.

Schmidt A. L., Zollo, F, Scala A, Betsch C, and Quattrociocchi W. (2018). 'Polarization of the vaccination debate on Facebook.' *Vaccine* 36(25), 3606-3612.

Short J, Williams E and Christie B. (1976). *The Social Psychology of Telecommunication*. John Wiley: New York.

Sia C.L., Tan B.C.Y. and Wei K.K. (2002). 'Group Polarization and Computer-Mediated Communication: Effects of Communication Cues, Social Presence, and Anonymity.' *Information Systems Research* 13(1), 70-90.

Seibold D. R., Canary D. J. and Tanita-Ratledge N. (1983). 'Argument and group decision-making: Interim report on a structurational research program.' Paper presented at the annual meeting of the *Speech Communication Association*, Washington, DC.

Siebold D.R. and Meyers R.A. (2007). ‘Group Argument: A Structuration Perspective and Research Program.’ *Small Group Research* 38(3), 312-36.

Siegel J, Dubrovsky V, Kiesler S. B. and McGuire T.W. (1986). ‘Group processes in computer-mediated communication.’ *Organic Behavior and Human Decision Processes* 37(2), 157–187

Singer D.J., Bramson A, Grim P, Holman B, Jung J, Kovaka K, and Ranginani A. (2019). ‘Rational social and political polarization.’ *Philosophical Studies* 176, 2243-67.

Sunstein C. (2002). ‘The Law of Group Polarization.’ *The Journal of Political Philosophy* 10(2), 175-95.

Tollefsen D. (2007). ‘Group testimony.’ *Social Epistemology* 21(3), 299-311.

Turner J. C. (1985). ‘Social categorization and the self-concept: A social cognitive theory of group behaviour.’ In E. J. Lawler (Ed.), *Advances in Group Processes: Theory and Research* vol. 2. JAI Press: Greenwich, CT.

Utz S, Kerkhof P, and van den Bos J. (2011). ‘Consumers rule: How consumer reviews influence perceived trustworthiness of online stores.’ *Electronic Commerce Research and Applications* 11, 49-58.

Valacich J. S., Dennis A. R., Connolly T. (1994). ‘Idea generation in computer-based groups: A new ending to an old story.’ *Organic Behavior and Human Decision Processes* 57(3), 448–467.

Walther J. B. (2011). ‘Theories of computer-mediated communication and interpersonal relations.’ *The handbook of interpersonal communication* 4, 443-479.

Walther J. B., and Parks M. R. (2002). 'Cues filtered out, cues filtered in: Computer-mediated communication and relationships.' In Knapp M. L. and Daly J. A. (eds.), *Handbook of Interpersonal Communication*, 529– 563. Sage: Thousand Oaks, CA.

Willemsen L.M., Neijens P.C., and Bronner F. (2012). 'The Ironic Effect of Source Identification on the Perceived Credibility of Online Product Reviewers.' *Journal of Computer-Mediated Communication* 18, 16–31.

Yardi S and Boyd D. (2010). 'Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter.' *Bulletin of Science, Technology & Society* 30(5), 316-27.