Conclusions and Reflections

Tanzhe Tang

THESIS CHAPTER

Chapter 6 (pp. 155-176) of the doctoral dissertation *On the Methods for Explaining Polarization of Private and Unobservable Opinions: An Opinion-Behavior Co-Evolutionary Approach*, Delft University of Technology, 2022. This standalone file reproduces the final published chapter and adds only this bibliographic cover page.

ABSTRACT

This chapter concludes the efforts made in the thesis to achieve its research goals, with a special focus on the relations between the proposed methods. It identifies promising directions for future work, discusses the societal relevance of the thesis, and reflects on the field of opinion dynamics and polarization.

RECOMMENDED CITATION

Tang, T. (2022). Conclusions and reflections. In T. Tang, *On the Methods for Explaining Polarization of Private and Unobservable Opinions: An Opinion-Behavior Co-Evolutionary Approach* (Chapter 6, pp. 155-176). Delft University of Technology.

SOURCE

Final published dissertation:
https://doi.org/10.4233/uuid:b4b02f3b-b622-4226-8f8c-be26e2c52428
License: CC BY

6

CONCLUSIONS AND REFLECTIONS

Opinion polarization is a widespread and potentially dangerous social phenomenon that is now reshaping economies and societies around the world. This statement, which is now part of common sense, has been repeatedly endorsed by real-life events of multiple scales. The current riots between pro- and anti-vaccine camps provide the most recent worldwide endorsement, underlining the urgency of countering opinion polarization, especially in crises such as the COVID-19 pandemic.

Battling against opinion polarization requires scholarly efforts to help us understand what polarization is and what causes polarization. Echoing this requirement, the thesis aims to make methodological contributions to opinion polarization studies from a rarely explored perspective, which centers on the "private and unobservable" feature of opinions. The main idea is that opinions cannot be directly observed by others but need opinion-driven behaviors to mediate their interactions with the opinions of others. In order to help modelers put this idea into scientific practice, the thesis proposes four research methods that cover different parts of opinion polarization, with the ultimate goal of arriving at a more specific and realistic understanding of this mysterious phenomenon.

This chapter concludes the thesis by summarizing and discussing these methods. Section 6.1 will provide an overview of each method while highlighting their functionalities and positions in the literature. Section 6.2 will unveil the relations between the four methods, and present them as an integrated toolkit. Avenues for future studies will be discussed in Section 6.3, followed by the societal relevance in Section 6.4. Finally, Section 6.5 reflects on various issues that emerge from the research process, including my thoughts on the field of opinion dynamics/ polarization.

6.1. OVERVIEW OF THE METHODS PRESENTED IN THE THESIS

In Chapter 1, I have declared that the primary research goal of the thesis is "to develop methods that could systematically support the exploration of mechanisms that can explain the polarization of private and unobservable opinions", which can be further divided into four sub-goals. From Chapter 2 to 5, each chapter aims to achieve one of the sub-

goals by presenting a research method for opinion polarization studies. The relations between the sub-goals, chapters, and methods are summarized as follows:

- Sub-goal 1: To develop a formal and broadly applicable polarization measurement that is coherent with the notion of group.
  - corresponding chapter: Chapter 2.
  - corresponding method: Equal Size Binary Grouping Measurement (ESBGM).
- Sub-goal 2: To develop a modeling method that can incorporate the "private and unobservable" feature of opinions in opinion polarization mechanisms.
  - corresponding chapter: Chapter 3.
  - corresponding method: Action-Opinion Inference Model (AOI model).
- Sub-goal 3: To develop a modeling method to study the effect of the deviation from the "default strategy", such as obfuscation, on opinion polarization.
  - corresponding chapter: Chapter 4.
  - corresponding method: Obfuscation-based opinion dynamics model (obfuscation model).
- Sub-goal 4: To develop a unifying framework of the co-evolution of opinion and behavior to organize existing efforts and facilitate future works of opinion polarization mechanisms.
  - corresponding chapter: Chapter 5.
  - corresponding method: Framework of Modeling Opinion-Behavior Co-evolution for Explaining Opinion Polarization (MOBEP framework).

I opened the entire thesis by promising to fulfill these sub-goals (and by doing so, I also consequently fulfill the primary research goal) in Chapter 1, and now I can claim that the promise has been kept. In the rest of this section, I will briefly revisit these methods, together with their functionalities and positions in the literature.

6.1.1. ESBGM: A GROUP-BASED POLARIZATION MEASUREMENT

Chapter 2 presents a group-based polarization measurement derived from my newly proposed, clustering-inspired grouping method called Equal Size Binary Grouping (ESBG). The measurement is not named in Chapter 2 but is later called Equal Size Binary Grouping Measurement (ESBGM) in Chapter 5. ESBG divides the data set to be measured into two equal-size groups based on the similarities between data points, and ESBGM is built upon this group structure as a function that increases with between-group heterogeneity and decreases with within-group heterogeneity. By providing an elegant and practical manner to identify and utilize group structures, ESBGM solves the longstanding problem in polarization measurements that a proper group structure is absent. Apart from its relations with between- and within-group heterogeneity, ESBGM possesses a number of

other ideal properties, including continuity (ESBGM is a continuous function), dimensionality (ESBGM is able to deal with both uni- and multi-dimensional data), maximum & minimum (ESBGM is maximized when data points in the same group are identical, and the difference between the two groups is maximized; ESBGM is minimized when all data points are identical), and normalization (the value of ESBGM is in the range between 0 and 1).

In Chapter 2, ESBGM is compared with measurements of bimodality (e.g., Ashman's D and bimodal separation) and bipolarization (e.g., Wolfson's index). All these measurements rely on the bi-division of the data set, but they are not measuring the same thing. First, bimodality may be a necessary but not sufficient condition for polarization. Therefore, bimodality measurements only capture one of the many aspects of polarization measured by ESBGM. Second, ESBGM and bipolarization measurements represent two different lines of polarization measurement, namely the "Esteban & Ray's line" (which focuses on clusters) and "Wolfson's line" (which focuses on the disappearance of the middle class). In Chapter 5, I further compare ESBGM with the FM index – a representative polarization measurement that has been widely used in opinion dynamics studies. I find that ESBGM is more sensitive to the distribution of group sizes than the FM index, while the FM index is more sensitive to extreme opinions.

Each polarization measurement reflects a unique understanding of polarization, and I cannot judge if a measurement "universally" outperforms another. Therefore, instead of claiming ESBGM to be the best measurement, I argue that it uniquely captures the widely accepted notion that "group" is a crucial ingredient of polarization. In practice, I would suggest using multiple measurements in opinion polarization studies to obtain a broader view of the system of interest (see the case study in Chapter 5 for an example).

6.1.2. AOI MODEL: AN AGENT-BASED MODELING METHOD FOR CO-EVOLUTION OF BEHAVIOR AND OPINION

The second sub-goal asks for "a modeling method that can incorporate the private and unobservable feature of opinions in opinion polarization mechanisms". In other words, the modeling method should be able to "incorporate behaviors in opinion polarization mechanisms as a messenger of opinions". This method is then formally introduced in Chapter 3 in the form of an agent-based model called Action-Opinion Inference (AOI) model.

The AOI model is not the first opinion dynamics model to include both opinion and behavior in the same dynamics. In Chapter 3, it is compared with the CODA model. The notable difference is that opinions are continuous in the CODA model but are discrete in the AOI model. The discreteness in the AOI model makes it possible to use the deontic logic (i.e., permission, obligation, and prohibition) to describe various behavior-opinion relations, while in the CODA model, the relations between continuous behavior and discrete opinion are rather straightforward, and in some cases, oversimplified (see Section 5.2.2 of Chapter 5 for details). In Chapter 5, the MOBEP framework helps us systematically compare the AOI model with other behavior-opinion co-evolution models, especially the conformity model (Buechel et al., 2015) and the SNOAEs model (Zhan et al., 2021). Among all these models, the AOI model, as the name indicates, almost exclusively describes the behavior-opinion inference process (meaning "inferring one's opinion from her behav-

ior") in a Bayesian manner, whilst other models would simply ignore the complexity of the process by treating behaviors as "extreme" opinions[1].

6.1.3. OBFUSCATION-BASED OPINION DYNAMICS MODEL: AN AGENT-BASED MODELING METHOD FOR OBFUSCATION-BASED OPINION DYNAMICS

Obfuscation brings additional uncertainty to the dynamics of opinions, and thereby has a potentially significant role in opinion polarization. On the basis of the AOI model, Chapter 4 presents a modeling method for studying obfuscation in the form of an agent-based model. This method models obfuscation as an agent's strategy to choose the behavior that gives out the least information about her opinion while ensuring that the chosen behavior is allowed. At the other extreme, the method is also capable of modeling transparent agents who want their opinions to be accurately known by others.

Obfuscation represents strategies of "intelligent agents" that deviate from the "default strategy" of "naive agents", which is exactly what the third sub-goal asks us to model. The agents described in the AOI model are considered naive as their behaviors are purely and passively determined by their opinions, while the agents modeled by this method have a higher level of intelligence in the sense that their behaviors are also directed by their individual goals (e.g., hiding their opinions or disclosing their opinions) and knowledge about other agents (e.g., how other agents would interpret their behaviors). As stated in Chapter 4, similar concepts include deception, strategic ambiguity, and intentional vagueness, whose roles in opinion polarization are also worth exploring. In light of this argument, I can frame the obfuscation model as an "intelligent" extension of the AOI model.

Obfuscation is also a "special characteristic" (Dong et al., 2018) of agents. Other special characteristics that have been extensively studied in opinion dynamics studies include skepticism (Tsang & Larson, 2014), zealotry (Waagen et al., 2015; Verma et al., 2014; Mobilia, 2003), and stubbornness (Yildiz et al., 2013). Obfuscation differs from them as it is exclusive to behavior-included opinion dynamics (i.e., opinion-behavior co-evolution): in the absence of behavior, there is no room for agents to obfuscate. From this perspective, this method can be viewed as a novel approach to studying special characteristics of agents in opinion dynamics.

6.1.4. MOBEP FRAMEWORK: A UNIFYING FRAMEWORK OF MODELING OPINION-BEHAVIOR CO-EVOLUTION FOR EXPLAINING OPINION POLARIZATION

Chapter 5 presents an opinion-behavior co-evolution framework MOBEP (**M**odeling **O**pinion-**B**ehavior co-evolution for **E**xplaining Opinion **P**olarization), which is particularly designed to explain opinion polarization via agent-based modeling. The framework is constructed by five key components that together capture all aspects of opinion-behavior co-evolution. The functionality of the framework is twofold. First, it provides an overarching architecture to organize existing efforts devoted to modeling opinion-behavior

[1] This is possible when an agent's opinion is set to be a continuous variable ranging from 0 to 1, and her behavior is set to be a binary variable that is either 0 or 1. As a result, behaviors can be directly added to and subtracted from opinions. See Section 5.2 of Chapter 5 for a review.

co-evolution, which would advance the accumulation of insights and modeling techniques. Second, future works of opinion polarization would benefit from the framework as these insights and techniques shall provide practical guidance for building new models, and the overarching architecture will help modelers better place their works in the vast literature of opinion polarization.

In the framework, classic models of opinion dynamics that exclude behaviors are represented by one of the key components – direct opinion influence. In this sense, my framework is of a higher level of granularity than frameworks of classic opinion dynamics such as the framework proposed by Coates et al. (2018), or, in other words, the classic ones are the subsets of mine. Meanwhile, the framework is specific to opinion dynamics, and is thereby of a lower granularity level than general frameworks of agent-based models (e.g., Masad & Kazil, 2015).

6.2. RELATIONS BETWEEN THE METHODS

An intrinsic feature of this paper-based thesis is that each content chapter is relatively independent, and all the methods presented in these chapters can stand on their own feet. However, the thesis is not a collection of poems where the only connection between each chapter is the same authorship. In fact, there is an underlying thread connecting each method to form one integrated toolkit to fulfill the primary research goal.

The methods presented in Chapter 3-5 are all built upon one central notion: behaviors are the messenger of opinions (see Section 5.2.3 of Chapter 5 for details)[2]. The AOI model lays down the foundation for modeling opinion dynamics under this notion, and the obfuscation model extends the boundaries by taking into consideration agents' "deviate" strategy of obfuscation. The MOBEP framework takes one step back from specific modeling tasks (e.g., the AOI model and the obfuscation model) to obtain a bird's eye view of the opinion-behavior co-evolution. Finally, the measurement ESBGM connects these three methods to opinion polarization by measuring the results generated by them, and forms a complete toolkit that is particularly designed for opinion polarization studies.

To see how these methods work as one integrated toolkit, one first needs to understand the research approach of a typical opinion polarization study in computational sociology. Usually, such a study starts with a core assumption that is potentially important for the emergence of opinion polarization. The core assumption is then implemented in an agent-based model, together with other model components. Running simulations of the model leads to simulation results of opinion distributions, from which conclusions about the assumption will be drawn. The four methods facilitate the research approach in the following ways:

- **Creating assumptions**: Considering the potentially crucial role of behavior in opinion polarization, it is promising to use the AOI and the obfuscation model as starting points to develop core assumptions, such as (if) obfuscation can foster polarization. At the same time, the framework could inspire the creation of both core assumptions and other model components with its accumulation of insights gained from previous works.

[2] This notion is a refined version of the central notion of the AOI model ("learning opinions by observing actions"). See Chapter 3.

- **Translating assumptions to agent-based model**: The AOI model and obfuscation model can be used to translate core assumptions that are related to both opinions and behaviors into agent-based model components. The framework provides an architecture to organize all model components (including the core assumption) by helping modelers identify the relations between them. In addition, such an architecture makes it easier to modify and then employ existing implementations of these model components from previous works: for example, the case study provided in Chapter 5 benefits from different models encompassed by the framework: the opinion-driven behavior change component is modified from the conformity model (Buechel et al., 2015), and the direct opinion influence part is modified from the rejection model (Flache & Macy, 2011).

- **Translating results to conclusions**: Translating the agent-based model to simulation results requires programming skills, while translating the simulation results to conclusions requires a proper polarization measurement. A polarization measurement not only summarizes the state of the whole system but also reflects our understanding of polarization. The measurement ESBGM presented in Chapter 2 echoes the widely agreed understanding that the notion of group, instead of individual, is the crucial ingredient of polarization, and hence is an ideal tool to derive conclusions from simulation results.

To conclude, the four methods are united as one toolkit that covers almost all the steps that are required to explain opinion polarization by agent-based modeling.

Another interesting perspective to understand the relations between these methods is to consider the framework as a configuration that accommodates the other three methods. The AOI model and the obfuscation model provide valuable insights for the opinion-driven behavior change component and the behavior-opinion inference component of the framework, and naturally, ESBGM relates to the component of polarization measurement (see Section 5.3 of Chapter 5 for the framework components). From this perspective, I can alternatively conclude that the entire thesis is about the MOBEP framework presented in Chapter 5, while Chapter 2-4 are preparing relevant ingredients.

6.3. AVENUES FOR FUTURE RESEARCH

As far as I can imagine, future research based on this thesis contains at least three possible categories: (i) improving the methods proposed in the thesis, and (ii) utilizing these methods for opinion polarization studies motivated by new research questions, and (iii) developing new research methods based on the proposed methods. Despite their different relations with the thesis, the three categories shall work towards the same goal: to deepen our understanding of the polarization of "private and unobservable opinions".

6.3.1. IMPROVING THE PROPOSED METHODS

As the output of my Ph.D. research, this thesis inevitably suffers from limitations of time and is thereby far from perfection. Consequently, there are many interesting and promising avenues for follow-up research that aims to improve the methods proposed in the thesis from different aspects. Basically, such avenues may take approaches such as

testing the methods with relaxed or modified assumptions, calibrating or validating the methods according to empirical data, generalizing the methods for broader applicability, and implementing the method in a more user-friendly manner. At the end of each relevant chapter, I have already discussed a number of avenues for future research, but more have been identified after the papers were published. For the reader's convenience, they are now summarized and briefly explained as follows:

AVENUES FOR IMPROVING ESBGM

(1) Systematically comparing ESBGM with existing polarization measurements, especially in the context of various attribute distributions: In Chapter 2, I have already compared ESBGM with bimodality and bipolarization measurements, and in Chapter 5, ESBGM is compared with the FM index. In addition, there are many other widely used polarization measurements, such as the Esteban-Ray index and the SSG index, whose relations with ESBGM remain unknown. Comparing ESBGM with these measurements would help us discover its uniqueness in the vast literature, and find out in which situations ESBGM would be more suitable than its competitors. To make meaningful comparisons, it is recommended that apart from theoretical analysis (e.g., checking whether some properties, such as increased spread, are satisfied. See Section 2.6 of Chapter 2 for an example), these measurements (including ESBGM) should be applied to various types of attribute distributions in order to compare each measurement's "characteristics", namely the particular aspect(s) of a distribution that the measurement values the most in assessing opinion polarization. For example, in Chapter 5, I find out that compared to the FM index, ESBGM places more weight on extreme values than group sizes. Distributions of two or more peaks are favorable candidates as they can be used to test how each measurement responds to intermediates state between consensus and polarization.

(2) Re-examining selected polarization studies by re-measuring their results using ESBGM: The choice of measurements reflects the authors' understanding of opinion polarization, and is able to substantially affect the final results. Section 5.4.4 of Chapter 5 has provided a vivid example where the polarization levels reported by ESBGM and the FM index can be extremely different when the opinion difference is large but the group size difference is small (see Figure 5.4 of Chapter 5). Therefore, it becomes pivotal to re-examine important conclusions derived in previous polarization studies by re-measuring corresponding results using ESBGM. Current conclusions may be rewritten and new insights can be developed if there exist significant differences between the ESBGM-measured results and the existing ones, implying that a widely accepted measurement (which represents a widely accepted understanding of polarization) is essential to reach consensus about polarization mechanisms. Meanwhile, the difference could also tell if ESBGM is suitable for the context of that particular study, and if necessary, one can improve ESBGM accordingly.

(3) Implementing ESBGM based on clustering techniques other than K-means: To implement ESBGM, one first needs to divide the data set into two groups of equal sizes (i.e., implementing ESBG). In Chapter 2, I have provided a demonstration where ESBG (and hence ESBGM) is implemented based on the most famous and simple clustering technique: K-means. According to Jain (2010) and HajKacem et al. (2019), thousands

of clustering techniques have been published in the last sixty-some years, and K-means is by no means the only one of them that can be adapted for ESBGM implementation. Each clustering method has its own advantage in discovering hidden grouping structures from the data, and ESBGM can exploit these advantages by employing various clustering methods as its basis. Here I present a list of research directions that are worth noticing in this regard:

(a) **Fuzzy-clustering-based ESBGM:** A large number of clustering techniques, including K-means, are the so-called "crisp clustering", meaning that each data point exclusively belongs to one of the groups (Xu & Wunsch, 2009). This "crisp" feature is sometimes harmful to the applicability of ESBGM because exclusive group membership means it will be impossible to divide a data set into two equal-size groups if its size is an odd number. As a result of this, it remains undecided how to deal with this type of data sets ("*For the sake of simplicity, we only discuss systems whose size is an even number*", Section 2.3.2 of Chapter 2). This problem can be solved by replacing crisp clustering (e.g., K-means) with fuzzy clustering, where each data point belongs to all (in our context, two) groups with a certain probability. In other words, each data point, instead of being assigned to one of the groups, now has a membership of a certain degree (Xu & Wunsch, 2009). Such a "soft assignment" (Jain, 2010) is able to address the issue of applicability because a data set of an odd-numbered size can now be divided into equal "shares" of data points by a fuzzy clustering technique. Moreover, fuzzy clustering can eliminate group membership transitions caused by imperceptible changes of a data point near the group boundary, which is at least an intellectually unfavorable circumstance for some scholars. Introduction to fuzzy clustering can be found in most textbooks on clustering or machine learning (e.g., Xu & Wunsch, 2009; Aggarwal & Reddy, 2014; Bishop, 2006; Müller & Guido, 2016), while a number of software tools are available to perform the task in various programming languages (e.g., *fclust* [3] for R, and *fuzzy-c-means* [4] for python). Therefore, implementing ESBGM based on fuzzy clustering should be technically feasible.

(b) **Outlier-robust-clustering-based ESBGM:** The last thing we want to encounter while using ESBGM is an outlier, which may dramatically change the division of groups and thereby erode our trust in the results. Even if the outlier is not caused by mistakes, its disproportional influence on the entire system's polarization level (measured by ESBGM) is rarely favored. The problem comes from K-means' strong sensitivity to outliers (Chawla & Gionis, 2013; Olukanmi & Twala, 2017), and the solution is, straightforwardly, to employ outlier-robust clustering techniques. Several modifications of K-means have been proposed to address the problem. K-medoids undermines the influence of outliers on clustering by enforcing each centroid to be fully overlapped with one of the objects within the group (Kaufman & Rousseeuw, 1990; Bishop, 2006). K-means-- (pronounced "k means minus minus") (Chawla & Gionis, 2013) and K-means # (pronounced "k means sharp") (Olukanmi & Twala, 2017) are able to detect outliers while preserving the advantages of K-

[3] https://cran.r-project.org/web/packages/fclust/index.html
[4] https://pypi.org/project/fuzzy-c-means/

means, allowing us to decide later if or how these outliers will be included in the future measuring process.

(c) **Large-scale-clustering-based ESBGM:** To save us from unnecessary complexity, the demonstration data sets used in Chapter 2 are of relatively small scales (less than 300 objects and two dimensions), and the measuring process is extremely fast. With the development of data science, scholars now need to deal with extra large-scale data that could easily take the original K-means a significant amount of time[5] to deal with, indicating that the K-means-based ESBGM would be even slower. To speed up the measuring process, one simply needs to adopt one of the improved versions of K-means (e.g., Pelleg & Moore, 1999; Stoffel & Belkoniene, 1999; Kanungo et al., 2002), or specific clustering techniques such as CURE (Guha, 1998)[6]. The cost of improved computational efficiency is the increased complexity of implementation, which can be minimized if the relevant algorithm is presented as a "black box", such as an R package or Python module, to scholars.

AVENUES FOR IMPROVING THE AOI MODEL

(1) Combining the model with homophily: The partner selection mechanism of the original AOI model in Chapter 3 is based on the voter model, which states that the focal agent would randomly choose a neighbor to interact with. As stated in Chapter 5, compared to random selection, homophily, the tendency that like-minded people are more likely to interact with each other, is a more realistic assumption of the partner selection mechanism, and in many models plays a crucial role in determining the dynamics of the system (e.g., Mäs & Flache, 2013). Combining the AOI model with homophily should lead to significant improvements in its ability to model real-life situations.

(2) Testing the AOI model in various network structures other than Von Neumann neighborhood: The original AOI model runs on the Von Neumann neighborhood for two reasons: (i) to minimize the complexity caused by networks, and (ii) to maximize the clarity of visualization. Future studies can extend the applicability of the model by testing it in other networks such as small-world networks (which mimic real-life connections between acquaintances) and multiplex networks (for situations such as co-evolution of online and offline opinion dynamics).

(3) Analytically solving the model of general action-opinion relations: In Chapter 3, the model has been analytically solved under the simplest action-opinion relation. Despite the possible difficulty, analytically solving the model of more general relations could validate the simulation results and help us better understand the dynamics.

(4) Empirically validating the model at both micro and macro levels: Like most opinion dynamics models, the AOI model has not been empirically validated, which means its model setups (micro-level) and aggregated outcomes (macro-level) have not been compared with real data. In the field of opinion dynamics, it is a common practice to build a

[5] The time complexity of K-means is $O(Nkd)$, where N is the data size, and d is the number of data dimensions (Xu & Wunsch, 2009).

[6] At this moment I don't know if these techniques can be modified according to the requirements of ESBGM.

model purely upon theories and assumptions, but recently, a trend of empirical validation has emerged, aiming to improve the real-life reliability of the in-silico outcomes (see Section 6.5.4 for a detailed discussion). The trend implies that empirical validation should be a necessary step to improve and promote the AOI model if I would like it to be used by more scholars.

(5) Generalizing the model by adopting continuous numerical values to represent action-opinion relations: The AOI model uses matrices with deontic logic to describe action-opinion relations. An entry of the matrix, which represents the evaluation of an action by an opinion, must be either + (obligation), − (prohibition), or 0 (permission). This setup remarkably simplifies the model by dissolving the opinion spectrum into three deontic symbols, but makes it difficult to disentangle similar opinions or preferences of (slightly) different degrees. For example, between strict vegetarians and carnivores, there are many people having different levels of preferences for meat consumption. Because these people are allowed (by their opinions) to consume both meat and vegetable[7], all of them are generally modeled in the AOI model as "omnivores" that have equal preferences for both options. Choice models have provided us with a natural solution to this issue, which is to replace the discrete deontic symbols with continuous numerical values (see Chorus et al. (2021) for an example). For example, an opinion may be described by "70% preference for meat and 30% preference for vegetable", which better suits the reality where most people are located somewhere in the opinion spectrum between the extremes. To summarize, in this avenue, I will generalize the applicability of the AOI model by endowing it with the ability to model continuous opinions.

AVENUES FOR IMPROVING THE OBFUSCATION MODEL

The obfuscation model can be regarded as an extension of the AOI model, so all the avenues for improving the AOI model are also applicable here. At the end of Chapter 4, I have explicitly mentioned the following avenues: (i) modifying model assumptions such as single-layer/ undirected networks (to multi-layer/ directed networks), positive social influence (to the mixture of positive and negative influence), and sequential updating (to parallel updating), and (ii) calibrating the model to empirical data of public opinions instead of providing illustrative examples.

AVENUES FOR IMPROVING MOBEP FRAMEWORK

At the end of Chapter 5, two avenues for future research have been proposed to improve the usability of the framework: (i) developing libraries for its components, and (ii) implementing the framework in a web-based application.

6.3.2. UTILIZING THE PROPOSED METHODS

The value of a method can be appreciated only when being used. The examples in Chapter 4 and 5 provide demonstrations of using these methods in opinion polarization studies. An intuitive and potentially fruitful starting point for such studies is to transform classic models where opinions are "what social influence influences" to models where opinions

[7] Here I have adopted a similar setting as in Section 4.4.1 of Chapter 4, where everyone only has two options: either meat or vegetable.

are private and unobservable, serving as a shortcut to constructing the initial and basic knowledge structure for this inadequately explored field. The transferring will benefit from the methods presented in this thesis, and I also expect that a cornucopia of new methods and tools will be generated during the process.

Transforming opinions from visible spins to unobservable attributes of agents is a significant step in approaching a more realistic and specific understanding of opinion polarization, but it is neither the only nor the last step as there are other features that distinguish opinions in models and in real life. For example, when modeling the dynamics of an individual's opinion, scholars usually focus on the influences from other people (i.e., social influence), but constantly ignore the effects of the individual's own attributes, such as his/her personality, experience, memory, and ideology. The obfuscation model is an ideal stepping stone towards investigating this topic as it provides an elegant way to incorporate agents' non-opinion attributes (here: agents' strategy of disclosing opinions) in opinion dynamics.

Meanwhile, as a complex social process, the dynamics of a single opinion dimension (i.e., opinion of a single topic) is hardly a closed system but usually intertwined with others. This requires us to carefully model the interplay between opinion dimensions that usually go hand-in-hand, such as abortion and gun rights (Johnson, 1997), or COVID-19 disbelief and anti-vaccine (Bok et al., 2021). Unlike models of continuous (e.g., Martins, 2008) or binary opinions (e.g., Banisch & Olbrich, 2019), the AOI model is able to model the co-evolution of relevant opinions dimensions, especially when the deontic symbols are replaced with numerical values (see Section 6.3.1). At the same time, ESBGM can provide reliable assessments of the results given that it applies to multi-dimensional data.

Furthermore, the rapid development and widespread use of information technology – including artificial intelligence, big data, and machine learning – are substantially reshaping the dynamics of opinions, which means our opinions also receive influences from external factors, such as social bots, recommendation systems, and online personalization that are driven by these new technologies. Investigating opinion dynamics in the era of information technology has obtained remarkable attention (e.g., Dandekar et al., 2013; Mäs & Bischofberger, 2015; Stella et al., 2018; Ross et al., 2019; Perra & Rocha, 2019; Keijzer & Mäs, 2021), but additional efforts are still in need to keep pace with new realities and trends. These technologies also need "behaviors" to influence opinion dynamics. For instance, social bots will mimic human behaviors like replying to messages, and recommendation systems will display personalized information to their users. Therefore, the topic falls within the scope of "opinion-behavior co-evolution", where the MOBEP framework can help. In the case of social bots, one can consider both social bots and human users as agents, and assume that social bots have fixed opinions while human users have changeable opinions that are open to social influences. The behaviors of agents are not only the messengers of opinions but also indicators of their identities (social bots or humans). The MOBEP framework can help translate these assumptions into opinion dynamics models and investigate the influence of social bots on opinion polarization.

6.3.3. DEVELOPING NEW METHODS: WITH AN EXAMPLE OF NETWORK-BASED POLARIZATION MEASUREMENT

Opinion polarization is an enormous topic with dark corners that need to be enlightened. The four methods presented in the thesis shed light on some of them, but there are more to be explored. Developing new methods based on these proposed methods is a shortcut to further explorations, as the latter have provided solid theoretical bases that promise a relatively high success rate. Different from the improved versions (See Section 6.3.1), new methods are supposed to be less directly connected with the original ones, whose core ideas, rather than their technique details, should be valued most. In the rest of this subsection, I will show one example of such new methods: a conflict-oriented network polarization measurement inspired by ESBGM.

New methods shall approach currently unanswerable questions. In the domain of polarization measurement, probably the most fundamental question is how to incorporate the "conflict nature" of polarization, which most likely originates from Esteban and Ray's claim that polarization is a "relevant correlate of potential or open social conflicts" (Esteban & Ray, 1994). Given this nature, measuring polarization becomes meaningless if no connections (and hence no conflicts) exist between people (Guerra et al., 2013). Distribution-based measurements (including all the measurements in Chapter 2) can stay away from this issue because they don't feature any network strictures (alternatively, we can say they implicitly assume that everyone is connected with everyone else), which leaves us with network-based measurements as the only option to incorporate this nature. The problem is, in network-based measurements, network edges usually don't represent connections ("I know you") but indicate endorsement ("I like you/ your opinion"), making it also incompatible with the "conflict nature".

In short, the question can be answered by a network-based polarization measurement that satisfies the following two properties:

(1) Interpersonal edges represent connections rather than endorsement; and

(2) The heterogeneity (in terms of the attribute(s) of interest) between disconnected people doesn't contribute to the polarization level.

Concerning the second property, it is natural to ask, if two previously disconnected people are now connected, will polarization go up? If we consider the connection as a channel where conflicts occur, then an additional connection between people of different groups (here: endogenously emerging groups, see Section 2.2 of Chapter 2 for details), which is an extra between-group edge in the network, shall increase polarization since it gives two enemies the chance to fight. If the new connection is between members of the same group (i.e., a new within-group edge), the story is more complex. One may argue that the new connection creates the possibility of cooperation to fight the other group, and polarization shall decrease (hereafter referred to as opinion Approach *A*). Others may argue that the new connection also offers the possibility of potential conflicts between the same-group members as long as they are not identical, which indicates that polarization should go up (hereafter referred to as Approach *B*).

ESBGM inspires us to deal with both approaches. By applying ESBG based on crisp clustering (such as K-means) to the data without considering its network structure, we

should obtain two groups of equal sizes. Different from ESBGM, now only the heterogeneity between connected pairs of data objects from the same (different) group(s) will be added to the total within- (between-) group heterogeneity. In this way, if the measurement decreases (increases) with within- (between-) group heterogeneity, additional within- (between-) group edges will always decrease (increase) polarization because they increase the within- (between-) group heterogeneity, and this measurement should then meet the requirement of Approach *A*. Meanwhile, by applying ESBG based on fuzzy clustering (see Section 6.3.1), each data point (i.e., nodes) i ($i = 1, 2, ..., N$, where N is the size of the population) should be given a clustering coefficient p_{ik} ($k = 1, 2$), which is the possibility that i is in group C_k. Given all p_{ik} for all i and k, we can generate a collection of all possible partitions of the data sets (together with their occurrence rate) that are in line with the clustering coefficients. To give the simplest possible example, if we have two nodes $i = 1, 2$, and $p_{11} = 0.5$ and $p_{21} = 1$, then in order to preserve these coefficients, we need at least two partitions with equal occurrence rate of 0.5: (i) both nodes are in C_1, and (ii) node 1 in C_2 and node i in C_1. The measurement should be the average between-group heterogeneity of all possible partitions weighted by their occurrence rate. Therefore, all edges (connections) contribute positively to polarization, but those who are considered as within-group edges in most occurrences (which means they are most likely to be classified as within-group edges in Approach *A*) will contribute less than those being considered as between-group connections in most occurrences (which means they are most likely to be classified as between-group edges in Approach *B*), and hence this measurement will be suitable for Approach *B*.

This ESBGM-inspired network-based measurement is far from completion, but it shows how the methods proposed in the thesis can inspire new methods for new research questions. After all, the development of the four proposed methods also benefits from existing ones like the K-means clustering and the voter model, and only by this type of "innovation chain" can the frontiers of the field be extended.

6.4. SOCIETAL RELEVANCE

In the introductory chapter, I have commented on the King's speech by pointing out that polarization is more than a tone of speech but a dreadful reality. Although I do not fully agree with His Majesty's opinion on polarization, I do utterly support his assertion that "[i]f every time is a time of transition, social change is a constant."[8] The year of 2021 acts as a live proof of this assertion: at the time of writing Chapter 1 (which was around mid-September), the number of daily COVID-19 cases in the Netherlands had been well maintained at a relatively low level due to the growing vaccination coverage and possibly the implementation of curfews and lockdowns. A few months later, when people were preparing for a cheering holiday season (compared to 2020), a possibly more infectious variant – Omicron – started to burst into prominence, making most optimistic predictions awfully hilarious. Opinions are even more mysterious than viruses, as there is neither a

[8] This is the official translation of the following (in Dutch): "*Als elke tijd overgangstijd is, is maatschappelijke verandering een constante.*" The full speech is available at https://www.koninklijkhuis.nl/documenten/toespraken/2021/09/21/troonrede-2021 (Dutch) and https://www.royal-house.nl/documents/speeches/2021/09/21/speech-from-the-throne-2021 (English, where the quotation in the main text is taken).

PCR test to identify one's hidden opinion, nor a dose of vaccine to slow down the spread of harmful ideas. Winston Churchill, who lost the 1945 British election as a war hero, and Hillary Clinton, who was expected to become the first female President of the United States by most polls in 2016, should agree with us with no hesitation.

So, if our societies, where opinions play a key role, simply hate pre-written scripts, why bother modeling them in the first place? It turns out that "all models (of opinion dynamics) are wrong, but some are useful"[9] in the sense that models developed in the ivory tower can actually help mitigate some problems – such as opinion polarization – that emerge during social changes, and the methods provided in this thesis could make this process easier.

In November 2021, YouTube announced that dislike counts would become invisible to viewers with the excuse that this change would "create an inclusive and respectful environment"[10] for creators. This announcement had caused widespread anger from both creators and viewers, questioning YouTube's intention and worrying about its consequence, such as whether the removal of dislike counts will alleviate or aggravate online opinion polarization. Luckily the methods of this thesis can help solve this problem as we can frame it as the opinion-behavior dynamics since most viewers' opinions are expressed by clicking the like or dislike bottom. Therefore, it will be very promising to use the AOI model, the MOBEP framework, and the polarization measurement ESBGM to study the opinion dynamics and its resulting polarization level before and after the removal of dislike counts. Furthermore, the removal is indeed an example of obfuscating behaviors, so the obfuscation model will also help capture the subtle changes in the minds and actions of both viewers and creators initiated by this controversial change. Although I don't have the space or time to include such a study in this very last chapter, the above description is already a good advertisement for using the methods of the thesis to mitigate real-life opinion polarization issues.

Over the past decades, battling opinion polarization remains a cheap talk or easy promise from politicians and business magnates who may actually benefit from opinion polarization. With research methods like those presented here, every stakeholder – from governments to ordinary citizens – can find their places in the battle as anything that is potentially relevant to opinion polarization can be formally analyzed (by scholars using these methods) and thereby practical suggestions – such as "establishing/ abolishing Internet real-name system" for policy makers, "keeping/ removing the dislike count" for social media companies, and "allowing/ rejecting cookies" for Internet users – will then be given. After all, opinions are private and unobservable, so everyone should be responsible for his/her own opinions.

6.5. REFLECTIONS

This section introduces my reflections on the research of this thesis. They are not directly related to my research goals, but may inspire and guide future studies of opinion dynamics and polarization.

[9] The quotation "all models are wrong, but some are useful" is a popular aphorism among modelers, which can be traced back to George Box, see https://www.lacan.upc.edu/admoreWeb/2018/05/all-models-are-wrong-but-some-are-useful-george-e-p-box/.

[10] See https://blog.youtube/news-and-events/update-to-youtube/ (accessed on 4 December 2021).

6.5.1. OPINION DYNAMICS VERSUS OPINION POLARIZATION

If the reader is not familiar with the topic of the thesis, he/she may be confused about two terms that often go hand-in-hand, namely "opinion dynamics" and "opinion polarization". The relation or difference between these two terms may be the first thing in the field that many people always wanted to know but were afraid to ask. This situation becomes more confusing when these two terms are used interchangeably in many papers.

If we keep a distance from the mixed usages, the difference is actually quite clear. Simply from its name, we know that "opinion dynamics" refers to the "dynamics of opinions", or "changes in opinions". Du et al. (2017) give a more specific definition that opinion dynamics is "the process in which agents form and update their opinions over time". Besides, the term "opinion dynamics" also serves as the name of a huge collection of models that describe this process ("opinion dynamics models"), as well as the name of the field where these models live ("the field of opinion dynamics").

Opinion polarization is a property of opinion distributions[11], whose meaning may differ across studies (see Chapter 2 for a comprehensive review and discussion). In other words, opinion polarization characterizes the outcome of the opinion dynamics that happened before. Traditionally, scholars focus on three types of outcomes generated by opinion dynamics, namely consensus (everyone has the same opinion), diversity (a large number of different opinions coexist), and polarization. I argue that these three outcomes can be all united under the name of polarization: consensus is the state of minimum polarization, and diversity is associated with medium levels of polarization. I can thereby draw the conclusion that every opinion dynamics model is somehow related to opinion polarization since they will end up in either consensus (low polarization), diversity (medium polarization), or (high) polarization. As a result, the boundary between opinion dynamics studies and opinion polarization studies is blurred, and this could be one of the reasons that are responsible for readers' confusion.

The above reflection provides us with new thoughts about opinion polarization research. As argued above, every outcome of an opinion dynamics model is related to opinion polarization. While a model that generates opinion polarization tells us why/how it happens, a model that generates consensus or diversity tells us how it can be avoided, which is indeed what the public and policy makers are expecting from opinion polarization studies. In this sense, we should not limit ourselves to the traditional approach that focuses on the search for models that lead to opinion polarization. Instead, equal attention should be given to other models in order to study how to promote consensus and diversity.

6.5.2. NEED FOR RIGOROUS DEFINITIONS OF COMMONLY KNOWN CONCEPTS

Human languages are vague. Expressed in the same words, what different people really mean can be very different, and this is how misunderstanding, distrust, and conflict start. As scholars are also humans, it is not surprising to find out that just like everyone else, they may use the same term while referring to different things. A more common phenomenon is that the scholarly usage of a term is different from its everyday usage, keeping relevant studies in an ivory tower isolated from the general public and policy makers.

[11] There are studies that consider opinion polarization as a property of the process of opinion dynamics (Dandekar et al., 2013), but this type of view is less popular.

This thesis has witnessed such vagueness in opinion dynamics research, which in fact motivated the whole thesis (see Chapter 1 for details). The most obvious example is about the term "opinion". Flache et al. (2017) claim that "opinion" generically refers to "what social influence influences", including concepts that are different from "opinion" such as "belief", "behavior", and "attitude". Meanwhile, in the models where opinion dynamics is coupled with behaviors (e.g., Martins, 2008), "opinion" no longer represents "behavior" and requires a less generic definition. Outside the field, the general public may feel awkward when seeing opinions being calculated and analyzed as numerical variables and observable properties of agents. It turns out that there is no consensus regarding the meaning of "opinion" whether within academia or between academia and the general public, which would lead to a series of chaotic consequences such as the hampering of knowledge accumulation and the misunderstanding between scholars and the public. Similar examples can be found in studies of opinion polarization, where "polarization" is usually considered as common knowledge that requires no explicit explanation, although people actually have very different understandings of this term (see Chapter 2).

The problem described above is caused by the lack of rigorous definitions of these concepts. Unlike natural or mathematical science, social science usually concerns concepts that are commonly known in our everyday life, such as "opinion" and "polarization". Scholars are likely to overlook the need for rigorous definitions of these concepts because they would assume that readers already know them. However, readers and even the scholars themselves may only have unclear and different understandings of these concepts. As a result, everyone interprets the concepts in their own ways, and it becomes difficult to tell if studies that carry the same name are actually talking about the same thing, and if the public has been correctly informed. To solve this problem, I would like to call for scholarly attention to the need for rigorous definitions of commonly known concepts in social science research in order to formally translate concepts in daily life into concepts in scientific contexts.

6.5.3. COMPLEXITY AND SIMPLICITY IN OPINION DYNAMICS MODELS

All models are simplifications of reality, but some models are simpler than others. In this thesis, models of different levels of complexity have been proposed. The simplest one should be the AOI model presented in Chapter 3, whose core is no more than a generalized voter model. The most complex model lies in the case study of mask wearing presented in Chapter 5, which is a combination of the conformity model (Buechel et al., 2015), rejection model (Flache & Macy, 2011), and beta-binomial model (Murphy, 2012; Khalvati et al., 2016, 2019), with modifications inspired by many other studies. As a case study that aims to illustrate how the framework – the main contribution of this chapter – works, this model has already cut off many other possibilities of modeling (e.g., modeling informational influence by a positive social influence model instead of a rejection model), which further indicates that a formal opinion dynamics model can be even more complex than the most complex one in this thesis.

It does not surprise us that opinion dynamics models have various complexity levels, but I am curious if there is any guiding philosophy about choosing the complexity level for model design. This is a non-trivial issue as everyone loves simplicity while everyone also wants a model that is close to reality. A few decades ago when scholars first attempted

to model opinion dynamics, they had few materials (e.g., modeling techniques, existing models, theories, etc.) at hand, which forced them to come up with simple and elegant models such as the voter model and DeGroot model. At our current moment when the field has been far from its infancy, staying simple becomes more difficult. On the one hand, most of the basic topics have been explored, and the remaining research questions, such as opinion polarization, are not likely to be answered by simple models. On the other hand, existing literature has warned us that certain types of model setups (e.g., modeling opinion as an observable feature of agents, not considering social network effects) are oversimplified, and more complex replacements are needed. Meanwhile, nobody wants to read a complex model whose description may occupy four or five pages: when you have (painfully) arrived at page four, you may have already forgotten what's happening on page one. In addition, unless the model is carefully built upon empirical data, an unnecessarily complex model usually means narrow applicability whose conclusions can hardly be generalized. Therefore, for those who want to study opinion dynamics in a realistic but general (i.e., not limited to particular real-life conditions such as location, topic, etc.) manner, it is crucial to find a balance between complexity and simplicity.

Let's first take a look at the simplest model in the thesis – the AOI model, whose aim is to introduce the notion of "learning opinions by observing actions" to opinion dynamics instead of replicating real-life situations. Therefore, for the AOI model, the priority is to formally and theoretically model the microscopic mechanism of how opinions can be learned by observing actions (i.e., the action-opinion inference process), while other model components that are not directly related to the notion, such as network structures and dimensionality of opinions, are given much less attention: in the original model, I simply choose the Von Neumann neighborhood as the network (which is somewhat a default choice for agent-based/ opinion dynamics models), and intentionally ignore the empirical value of modeling multi-dimensional opinions (see Chapter 2 for a discussion about uni- and multi-dimensional opinion polarization). Doubtlessly, these model components are also generally important in modeling opinion dynamics, but they are not the main characters of the AOI model, and thereby choosing anything different from the default option (which is also the simplest in most cases) will unnecessarily complicate the model. However, one of the necessary and promising directions for future studies is to test the AOI model in different social networks and opinion dimensions to see if the conclusion obtained in the default setting (Von Neumann neighborhood and uni-dimensional opinion) is still valid. To summarize, the notion or concept of interest should be first introduced in the simplest possible manner, and more complexity shall be added.

On the contrary, the most complex model in the thesis – the mask-wearing model presented in Chapter 5 – specifies each model component carefully in order to approach the real situation. For example, the small-world network used in the model resembles the limited offline interactions between humans enduring lockdowns, and the co-existence of behavior dynamics and direct opinion dynamics resembles our hybrid life where online communications on social media and offline (non-verbal) interactions in local stores are mixed. To generalize, in a model whose aim is to give an accurate and detailed description of relevant social dynamics or phenomena, all model components are of equal importance as any unrealistic or oversimplified assumption about any component will harm the reliability of the model. Therefore, I cannot assign default options to the

model components as I did in the AOI model but need to carefully specify each of them based on either theories or real situations.

To conclude, I argue that the level of complexity of an opinion dynamics model should depend on its aim. If the model is used to introduce a new notion, concept, or assumption, other model components should be kept at a minimum level of complexity if they are not directly related to the implementation of the newly introduced subject. Examples include the AOI model in Chapter 3 and the obfuscation model in Chapter 4. If the model is used to describe the reality, all model components should be carefully specified, leading to a more complex but more realistic model such as the mask-wearing model presented in Chapter 5.

6.5.4. SOME THOUGHTS ON THE FIELD OF OPINION DYNAMICS

The field of opinion dynamics[12] is unique. You can find relevant research papers from both physics journals such as Europhysics Letters and sociology journals such as the Journal of Mathematical Sociology. From the physics side, opinion dynamics offers the opportunity to apply modeling techniques that were originally designed for particles and spins to human agents, which is a challenging but exciting journey with the expectation of witnessing the hidden similarity between physics systems and human societies. From the sociology side, opinion dynamics represents a new approach to understanding social phenomena, where sociologists can control the entire system as a simple and clean representation of our society and draw conclusions by simulating or analytically solving it. With the joint contribution from physicists and sociologists, the field has endured a rapid process of development in the last decades, resulting in a large and growing number of publications.

The increase in quantity does not necessarily indicate an improvement in quality. In general, there is a tendency that the most influential papers are always these decades-old ones. There is nothing wrong with an individual model that is built upon a classic masterpiece, but for a fast-growing field, this tendency, while endorsing these classic papers, leads to the question that if the development of the field has reached a plateau.

It may be premature to make any assertion at this moment, but undoubtedly new directions are in urgent need to drive the field out of the accumulation of models (instead of contributions) to the next stage. This thesis underlies one of the new directions that take into consideration the "private and unobservable" feature of opinions by introducing behaviors to opinion dynamics. This direction dives deeper into modeling opinion-opinion interactions between agents, while other directions may be devoted to different aspects or granularity levels of opinion dynamics. For example, one ongoing trend is to study the formation of opinions within each agent by breaking opinions into smaller components such as arguments (Mäs & Flache, 2013; Feliciani et al., 2021) and attitude elements (van der Maas et al., 2020). Furthermore, by integrating opinion formation with argumentation frameworks, we will be able to model deliberations inside an agent and between agents (Taillandier et al., 2019), implying that this approach is potentially useful in explaining opinion polarization in a more realistic and detailed manner (see Proietti (2017) for a minimalistic example).

[12] To avoid unnecessary complexity, here, "the field of opinion dynamics" also includes the field of opinion polarization. See Section 6.5.1 for a discussion.

Apart from these particular directions, the future of this interdisciplinary field is most likely to be driven by interdisciplinary forces. One of the driving forces comes from the techniques that originated from physics and mathematics, such as the hierarchical Ising model (van der Maas et al., 2020), random field Ising model (Tiwari et al., 2021), matrix-weighted Laplacian dynamics (Ahn et al., 2020), and sheaf model (Hansen & Ghrist, 2021). The adoption of these new techniques allows us to investigate a wide range of new topics, expanding the territory of opinion dynamics studies. The second driving force comes from empirical data. A common criticism of the field lies in the lack of empirical support. As a response, a growing number of opinion studies use empirical data, whether from surveys, field studies, or lab experiments, to calibrate and validate their models, or use opinion dynamics models (with simulations) to test, explain, and generalize the empirically observed findings (see Kozitsin (2021), and Keijzer & Mäs (2021) for some very recent examples). In the integration between opinion dynamics research, physics techniques, and empirical studies lies the next frontier of computational sociology, where the polarization of the "private and unobservable opinions" is likely to be better understood.

REFERENCES

[1] Aggarwal, C. C., & Reddy, C. K. (2014). Data clustering: Algorithms and applications. CRC Press.

[2] Ahn, H. S., Van Tran, Q., Trinh, M. H., Ye, M., Liu, J., & Moore, K. L. (2020). Opinion dynamics with cross-coupling topics: Modeling and analysis. IEEE Transactions on Computational Social Systems, 7(3), 632-647.

[3] Banisch, S., & Olbrich, E. (2019). Opinion polarization by learning from social feedback. The Journal of Mathematical Sociology, 43(2), 76-103.

[4] Bishop, C.M (2006).Pattern recognition and machine learning. Springer.

[5] Bok, S., Martin, D. E., & Lee, M. (2021). Validation of the COVID-19 Disbelief Scale: Conditional indirect effects of religiosity and COVID-19 fear on intent to vaccinate. Acta Psychologica, 219, 103382.

[6] Buechel, B., Hellmann, T., & Klößner, S. (2015). Opinion dynamics and wisdom under conformity. Journal of Economic Dynamics and Control, 52, 240-257.

[7] Chawla, S., & Gionis, A. (2013). k-means–: A unified approach to clustering and outlier detection. In Proceedings of the 2013 SIAM International Conference on Data Mining (pp. 189-197). Society for Industrial and Applied Mathematics.

[8] Chorus, C., van Cranenburgh, S., Daniel, A. M., Sandorf, E. D., Sobhani, A., & Szép, T. (2021). Obfuscation maximization-based decision-making: Theory, methodology and first empirical evidence. Mathematical Social Sciences, 109, 28-44.

[9] Coates, A., Han, L., & Kleerekoper, A. (2018). A unified framework for opinion dynamics. In Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (pp. 1079–1086).

[10] Dandekar, P., Goel, A., & Lee, D. T. (2013). Biased assimilation, homophily, and the dynamics of polarization. Proceedings of the National Academy of Sciences, 110(15), 5791-5796.

[11] Dong, Y., Zhan, M., Kou, G., Ding, Z., & Liang, H. (2018). A survey on the fusion process in opinion dynamics. Information Fusion, 43, 57-65.

[12] Du, E., Cai, X., Sun, Z., & Minsker, B. (2017). Exploring the role of social media and individual behaviors in flood evacuation processes: An agent-based modeling approach. Water Resources Research, 53(11), 9164-9180.

[13] Esteban, J. M., & Ray, D. (1994). On the measurement of polarization. Econometrica: Journal of the Econometric Society, 62(4), 819-851.

[14] Feliciani, T., Flache, A., & Mäs, M. (2021). Persuasion without polarization? Modelling persuasive argument communication in teams with strong faultlines. Computational and Mathematical Organization Theory, 27(1), 61-92.

[15] Flache, A., & Macy, M. W. (2011). Small worlds and cultural polarization. The Journal of Mathematical Sociology, 35(1-3), 146-176.

[16] Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., & Lorenz, J. (2017). Models of social influence: Towards the next frontiers. Journal of Artificial Societies and Social Simulation, 20(4).

[17] Guerra, P. C., Meira Jr, W., Cardie, C., & Kleinberg, R. (2013). A measure of polarization on social media networks based on community boundaries. In Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media (pp. 215 - 224).

[18] Guha, S., Rastogi, R., & Shim, K. (1998). CURE: An efficient clustering algorithm for large databases. In Proceedings of ACM SIGMOD International Conference on Management of Data (pp. 73 – 84).

[19] HajKacem, M.A.B., N'Cir, CE.B., & Essoussi, N. (2019). Overview of scalable partitional methods for big data clustering. In O. Nasraoui & CE. Ben N'Cir (Eds.), Clustering methods for big data analytics. Springer.

[20] Hansen, J., & Ghrist, R. (2021). Opinion dynamics on discourse sheaves. SIAM Journal on Applied Mathematics, 81(5), 2033-2060.

[21] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651-666.

[22] Johnson, N. J. (1997). Principles and Passions: The Intersection of Abortion and Gun Rights. Rutgers Law Review, 50, 97-197.

[23] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7), 881-892.

[24] Kaufman, L., & Rousseeuw, P. J. (1990). Finding groups in data: An introduction to cluster analysis. John Wiley & Sons.

[25] Keijzer, M. A., & Mäs, M. (2021). The strength of weak bots. Online Social Networks and Media, 21, 100106.

[26] Khalvati, K., Mirbagheri, S., Park, S. A., Drehere, J. C., & Rao, R. P. (2019). A Bayesian theory of conformity in collective decision making. In Proceedings of the 33rd International Conference on Neural Information Processing Systems (pp.9702–9711).

[27] Khalvati, K., Park, S. A., Dreher, J. C., & Rao, R. P. (2016). A probabilistic model of social decision making based on reward maximization. In Proceedings of the 30th International Conference on Neural Information Processing Systems (pp. 2901-2909).

[28] Kozitsin, I. V. (2021). Opinion dynamics of online social network users: A micro-level analysis. The Journal of Mathematical Sociology. https://doi.org/10.1080/0022250X.2021.1956917

[29] Martins, A. C. (2008). Continuous opinions and discrete actions in opinion dynamics problems. International Journal of Modern Physics C, 19(04), 617-624.

[30] Mäs, M., & Bischofberger, L. (2015). Will the personalization of online social networks foster opinion polarization? SSRN. http://ssrn.com/abstract=2553436

[31] Mäs, M., & Flache, A. (2013). Differentiation without distancing. Explaining bi-polarization of opinions without negative influence. PloS ONE, 8(11), e74516.

[32] Masad, D., & Kazil, J. (2015). MESA: An agent-based modeling framework. In Proceedings of the 14th Python in Science Conference (pp. 51-58).

[33] Mobilia, M. (2003). Does a single zealot affect an infinite group of voters?. Physical Review Letters, 91(2), 028701.

[34] Müller, A.C., & Guido, S. (2006). Introduction to machine learning with Python: A guide for data scientists. O'Reilly Media.

[35] Murphy, K. P. (2012). Machine learning: A probabilistic perspective. The MIT Press.

[36] Olukanmi, P. O., & Twala, B. (2017). K-means-sharp: Modified centroid update for outlier-robust k-means clustering. In 2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech) (pp. 14-19).

[37] Pelleg, D., & Moore, A. (1999). Accelerating exact k-means algorithms with geometric reasoning. In Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 277-281).

[38] Perra, N., & Rocha, L. E. C. (2019). Modelling opinion dynamics in the age of algorithmic personalisation. Scientific Report, 9(1), 7261.

[39] Proietti, C. (2017). The Dynamics of group polarization. In A. Baltag, J. Seligman & T. Yamada (Eds.), Logic, rationality, and interaction (LORI 2017 Proceedings) (pp. 195-208), Springer.

[40] Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G., & Stieglitz, S. (2019). Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. European Journal of Information Systems, 28(4), 394-412.

[41] Stella, M., Ferrara, E., & De Domenico, M. (2018). Bots increase exposure to negative and inflammatory content in online social systems. Proceedings of the National Academy of Sciences, 115(49), 12435-12440.

[42] Stoffel K., & Belkoniene A. (1999). Parallel k/h-Means clustering for large data sets. In Proceedings of Euro-Par'99 Parallel Processing (pp. 1451-1454). Springer.

[43] Taillandier, P., Salliou, N., & Thomopoulos, R. (2019). Coupling agent-based models and argumentation framework to simulate opinion dynamics: Application to vegetarian diet diffusion. Social Simulation Conference 2019. Mainz, Germany. https://doi.org/10.3929/ethz-b-000383850

[44] Tiwari, M., Yang, X., & Sen, S. (2021). Modeling the nonlinear effects of opinion kinematics in elections: A simple Ising model with random field based study. Physica A: Statistical Mechanics and its Applications, 582, 126287.

[45] Tsang, A., & Larson, K. (2014). Opinion dynamics of skeptical agents. In Proceedings of the 2014 International Conference on Autonomous Agents and Multiagent Systems (pp. 277-284).

[46] van der Maas, H. L. J., Dalege, J., & Waldorp, L. (2020). The polarization within and across individuals: The hierarchical Ising opinion model. Journal of Complex Networks, 8(2), cnaa010.

[47] Verma, G., Swami, A., & Chan, K. (2014). The impact of competing zealots on opinion dynamics. Physica A: Statistical Mechanics and its Applications, 395, 310-331.

[48] Waagen, A., Verma, G., Chan, K., Swami, A., & D'Souza, R. (2015). Effect of zealotry in high-dimensional opinion dynamics models. Physical Review E, 91(2), 022811.

[49] Xu, R., & Wunsch, D. C. (2009). Clustering. Wiley. https://doi.org/10.1002/9780470382776

[50] Yildiz, E., Ozdaglar, A., Acemoglu, D., Saberi, A., & Scaglione, A. (2013). Binary opinion dynamics with stubborn agents. ACM Transactions on Economics and Computation (TEAC), 1(4), 1-30.

[51] Zhan, M., Kou, G., Dong, Y., Chiclana, F., & Herrera-Viedma, E. (2021). Bounded confidence evolution of opinions and actions in social networks. IEEE Transactions on Cybernetics. https://ieeexplore.ieee.org/document/9325877