Social Polarization: Towards the Text Mining Study

Roman Kyrychenko –
grant funded researcher, Faculty of Social Sciences, University of Helsinki, Finland
E-mail: kirichenko17roman@gmail.com,
ORCID: https://orcid.org/0000-0002-1610-4352

Роман Кириченко –
запрошений дослідник,
факультет соціальних наук, факультет соціології, Університет Гельсінкі, Фінляндія

DOI: 10.29038/2306-3971-2022-01-17-24

This article suggests a way to concretize the concept of social polarization, which will be most suitable for a) empirical operationalization; b) a complete description of existing social conflicts; c) will be consistent with existing sociological theories. The implementation of this task opens the way to studying this phenomenon through the method of text mining. We see two main problems with the concept of social polarization in sociology: (1) Social polarization is used as a beautiful metaphor to describe contemporary political situations, not as a strong operationalized concept. The concept must create vast opportunities to study social reality, interpret more processes; (2) The mathematic interpretation of social polarization is conducted on somewhat idealized distributions; there is a lack of real empirical data verification. These two problems also create one big problem: mathematical conceptualization of social polarization and empirical studies of social polarization are unrelated. We propose a way to solve this problem through the construction of our social polarization theoretical framework. The way that allowed us to do this was in the concretization of social polarization and its connection with sociological theories of conflict. The article's key idea is to show that this concept is suitable for operationalization for two reasons: its ability to describe the causes and nature of social conflicts and its measurability.

This article also discusses the main modern social polarization theories, their features, advantages, and disadvantages. Since the concept of social polarization is mostly the focus of political science research, the author's goal was to find opportunities to use this concept in sociology and the ideas that will allow it. There are currently two approaches to studying this issue: the party association approach and the opinion-based group approach. An important task, which was also solved in this article, is the concept's connection with the sociological concepts of conflict. The path was found using the concept of Lipset and Rokkan. This concept's key advantage is the combination of social inequality, conflict, attitudes, and social distance. Typically, these concepts are used separately to explain social cleavages. The concept of polarization, in this case, allows them to be integrated into a single whole.

Кириченко Роман. Соціальна поляризація: назустріч вивченню тексту. У цій статті запропоновано спосіб конкретизації поняття соціальної поляризації, який буде найбільш придатним для: а) емпіричної операціоналізації; б) повного опису наявних соціальних конфліктів; в) відповідності соціологічним теоріям. Реалізація такого завдання відкриває шлях до вивчення цього явища методом інтелектуального аналізу тексту. Ми бачимо дві основні проблеми з концепцією соціальної поляризації в соціології: (1) соціальна поляризація використовується як красива метафора для опису сучасних політичних ситуацій, а не як сильна операціо-налізована концепція. Концепція повинна створити широкі можливості для вивчення соціальної реальності, інтерпретації більшої кількості процесів; (2) математичне тлумачення соціальної поляризації проводиться на дещо ідеалізованих розподілах; відсутня реальна перевірка емпіричних даних. Ці дві проблеми також створюють іншу велику проблему: математична концептуалізація соціальної поляризації та емпіричні дослідження

соціальної поляризації не пов'язані. Ми пропонуємо шлях розв'язання цієї проблеми шляхом побудову власної теоретичної основи соціальної поляризації через конкретизацію соціальної поляризації та її зв'язок із соціологічними теоріями конфлікту. Ключова ідея статті полягає в тому, щоб показати, що це поняття придатне для операціоналізації з двох причин: через його здатність описувати причини й природу соціальних конфліктів та його вимірність.

Ключові слова: соціальна поляризація, групи думок, теорія соціальної ідентичності, теорія розколів.

INTRODUCTION

The modern development of theories of social polarization is going through an essential stage of evolution. If earlier, society's polarization was associated mainly with ideological division or social inequality, more and more researchers would talk about the importance of affects in citizens' polarization in the last decade. Because of this, the term «social polarization» is becoming popular in the scientific literature.

We propose to use the term «social polarization» as one that more fully characterizes the existing social distinctions.

The study of social polarization is quite complex, primarily due to differences in understanding of this concept. The authors of the article «Disambiguation of social polarization concepts and measures» counted as many as 9, but in general, they can be aggregated into two groups:

The polarization meaning is more associated with socio-economic stratification (Andersen, 2004; Baum, 1997; Hamnett, 1994). At the level of empirical research, this approach often leads to a banal analysis of the polarization of incomes of different groups of the population, which, in our opinion, eliminates the explanatory possibilities of this concept.

The polarization meaning is more associated with differences in political views (Druckman, & Levendusky, 2019; Iyengar et al., 2019; Luttig, 2017; Mason, 2018). In this article, we focus on only the second meaning of this concept. In our opinion, this approach is more promising, as it allows us to explain through the concept of polarization the formation of different groups of identities in society, the conflicts between them.

To sum up, the first concept is more descriptive, while the second has excellent explanatory potential. However, this definition also does not satisfy us completely. The reason for this is its focus on describing political processes.

There are different names, such as political, ideological, party, and social polarization. In general, researchers attach the same meaning to them. The reason for this is that mostly these concepts are developed by political scientists, not sociologists. Because of this, the term «social polarization» is predominantly political. It is often explicitly understood as the United States' situation with strong opposition from supporters of the Democratic and Republican parties (Iyengar et al., 2012).

This situation creates a new **goal** for this publication - to highlight social polarization as primarily an object of sociological rather than political science interest.

1. THEORETICAL APPROACHES TO THE STUDY OF SOCIAL POLARIZATION: ANALYSIS AND DISCUSSION

1.1. Two Main Concepts of Social Polarization

The concept of social polarization has become popular since 2012 through Iyengar's publications, who suggested using the term to refer to differences between Democrats and Republicans. His rationale was that the differences between them were not so much an ideology, but in the persistent negative emotional attitude towards each other, strongly associated with the party (Iyengar et al., 2012). It has long been argued that ideology causes the most disagreements (Dalton, 2008). This approach substantially grew as an opposition to the theory of ideological polarization.

The theoretical basis of his concept Iyengar took the social psychological concept of Henry Tajfel's social identity. Tajfel's idea is that all people tend to categorize themselves and relate to a particular in-group. At the same time, this group becomes part of its identity, strengthened by building a stable border with out-groups (Tajfel et al., 1971). Iyengar imposed this on the association of US citizens with parties and studied differences of opinion between them. As we can safely say, most of the available studies of social polarization follow the Iyengar tradition. The following standard features characterize this tradition:

1) **Emphasis on Party Affiliation.** Party association is the basis for dividing people into separate poles. Instead, this is a practical reason – party associations are regularly studied in empirical sociological research,

and in most democracies, where sociological services operate, there is research data that researchers can use. Also, the political process helps to ensure that these associations are updated continuously due to the elections.

2) **Attention to the Formation of Emotional Attitudes, Socio-Psychological Aspects.** It was important not just to record the differences in views, but also the emotional attitude towards them. Willingness to hate is what allows authors who move within this tradition to declare the affectivity of polarization.

3) **Studying the Role of Social Media.** In his first publication on this topic, Iyengar paid very little attention to the media's role in social polarization. Another political scientist, Levendusky, began an ongoing study of their role within this tradition. He posed the following research question: «Do the media increase social polarization?». The study of media can be divided into three groups:

– study of the influence of party media;

– review of the power of mainstream media;

– examination of the impact of social media (Levendusky, & Malhotra, 2013).

There is a large group of publications with empirical studies of how the media will isolate polarized groups, particularly the formation of so-called «echo chambers». The main idea of these publications is that fragmentation of the information field makes it impossible to form consensus decisions.

However, this approach has its drawbacks. The first disadvantage is its orientation on political science. Research projects based on this concept depends on the party and focuses only on it. Moreover, it is almost impossible to use in countries with an unstable pariah system, where parties are continually changing, such as Ukraine. By that logic, social polarization cannot exist in such countries. Party affiliation leaves the development of this concept in the framework of political science analysis. Some researchers have adapted this scheme to other systems, for example, through equality between party affiliation and ideology (Lauka et al., 2018; Mason, 2015, 2018). Liliana Mason's theory of the identity-based approach to ideology, initiated by Levitin and Miller, is of particular interest. The essence of this approach boils down to the thesis that the critical element of ideology is not the system of values and views shared by their supporters, but the identity formed around the label of ideology. Because of this, Mason speaks of social ideological polarization, with parties only representing conservative and liberal ideological camps. Mason argues that social polarization is more not a difference in ideological positions but merely a rejection of another identity. Such a scheme provides more flexibility for use (for example, transfer to other political realities (Johnston, 2019; Tworzecki, 2019)), but it remains within the framework of political analysis.

The second disadvantage is Americanization. Even if we agree that social polarization exists only on party associations, this scheme is still more focused on American realities, where there is a transparent bipartisan system. This leads to the fact that social polarization remains a term to describe modern American politics' realities, rather than a universal sociological concept that can describe and explain a wide range of phenomena.

Sara B. Hobolt, Thomas J. Leeper, and James Tilley posed a research question: is social polarization limited by partisanship, which was the beginning of developing an alternative and more sociological understanding of this concept (Hobolt et al., 2020).

To do this, they turned to the theory of Lipset controversial topics that can provoke a different and sharp reaction in society. The result of this appeal was the development of the concept of aphetic polarization around opinion-based groups. The authors took the theory of social action Craig McGarty, Ana-Maria Bliuc, based on opinion-based groups. The latter turned to its development to give more flexibility to Teiffel's theory of social identities. They decided that group identity can be formed based on shared opinion (Bliuc et al., 2007; McGarty et al., 2009). The authors' idea is broader than traditional views on social polarization, which creates a broad sociological perspective of its study. From their point of view, partisanship is only one of many dimensions by which society can be divided into polarized groups. However, the division can be in another aspect, such as the attitude to a very socially relevant issue. Such a question in their study was Brexit.

Like the Iyengar tradition, their concept is primarily based on social identity theory and the division into in-group and out-group Tajfel and Turner (Tajfel et al., 1971).

The authors of the paper mentioned above identified three components of opinion-based groups of social polarization:

(1) in-group identification based on shared views

(2) differentiation of the in-group from the out-group that leads to in-group favorability and out-group denigration

(3) evaluative bias in perceptions of the world and in decision-making.

They also suggest that these three components be evaluated empirically. First of all, it is the detection of the presence of identification (mainly through self-reporting). In the second place - the opposition to the group of Others. The third is assessing how these groups differ from each other within the discourse on the question that polarizes them.

This scheme's great advantage is that it is entirely consistent with Iyengar's scheme does not deny it. Still, it only expands the possibilities of using the concept of social polarization in social research. It can be used, for example, to study the polarization in Ukraine around the 2019 presidential election. Partisan polarization has a significant advantage in permanence and constant actualization, but in many societies, the party is not as key an institution as in the United States.

Baldasarri's and Bearman's theory of political polarization stands out. They rely on their work on the theory of the interpersonal influence of Lazarsfeld and Berelson. The key idea of this theory is that micro-interactions accumulate and lead to macro-effects. Polarization, according to this theory, is a consequence of such micro-interactions of members of small groups. This theory underwent a significant revolution during the second half of the twentieth century because, until the 1990s, it was believed that the consequence of such interpersonal interaction is a common consensus. However, Friedkin and Johnsen developed a social influence network theory that showed how such networks could conflict. The authors used Fridkin's developments in group polarization modeling to model the change of views. Polarization increases when views in one part of society are radicalized as opposed to another. The model of change of opinion takes into account two factors:

– selection of interaction partners, which determines the overall structure of the discussion network;

– the process of interpersonal influence, which determines the dynamics of change of thought.
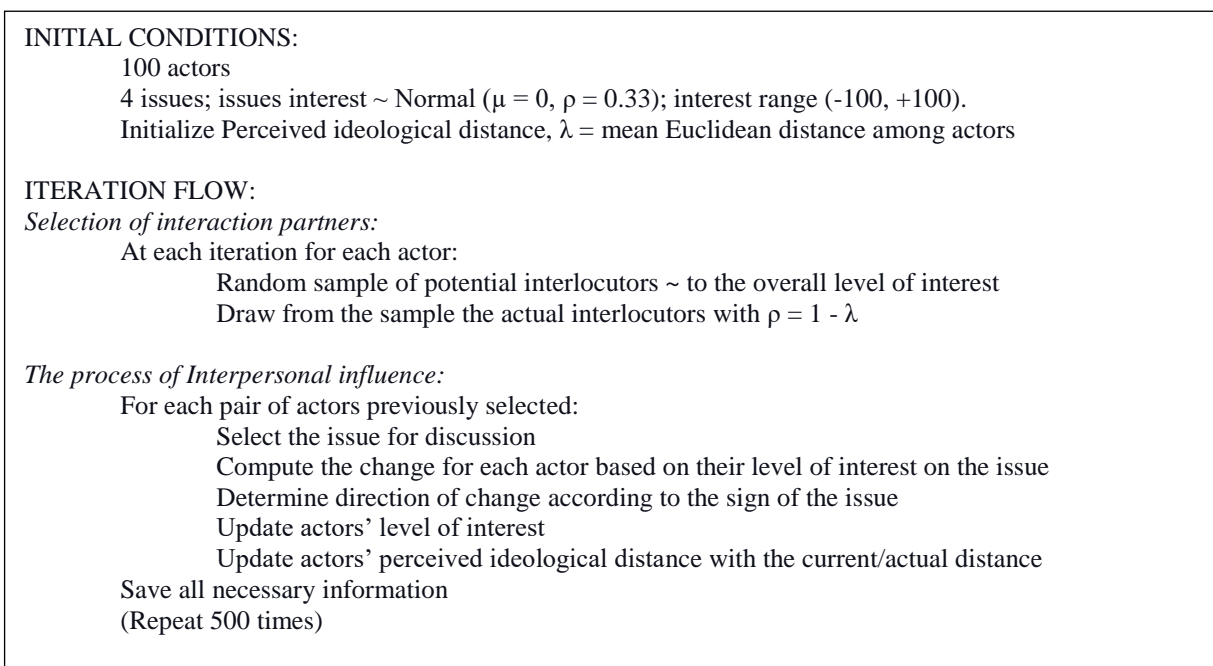
The model itself looks like this:

INITIAL CONDITIONS:
    100 actors
    4 issues; issues interest ~ Normal ($\mu = 0$, $\rho = 0.33$); interest range (-100, +100).
    Initialize Perceived ideological distance, λ = mean Euclidean distance among actors

ITERATION FLOW:
Selection of interaction partners:
    At each iteration for each actor:
        Random sample of potential interlocutors ~ to the overall level of interest
        Draw from the sample the actual interlocutors with $\rho = 1 - \lambda$

The process of Interpersonal influence:
    For each pair of actors previously selected:
        Select the issue for discussion
        Compute the change for each actor based on their level of interest on the issue
        Determine direction of change according to the sign of the issue
        Update actors' level of interest
        Update actors' perceived ideological distance with the current/actual distance
    Save all necessary information
    (Repeat 500 times)

Fig. 1. *Baldassari-Bearman Polarization Algorithm (Baldassarri, & Bearman, 2016)*

The simulations of the model show that with increasing iterations of interactions from the normal distribution of values, two polarities are distinguished.

Thus, from the point of view of the Baldassarri and Bearman model, the polarization follows from the structure of discussion networks. It is worth noting that the authors conducted simulations of the model, did not test them on real data. Therefore, there is a potential to use their theory for empirical verification (Baldassarri, & Bearman, 2016; McGarty et al., 2009).

Bearman and Baldassari's theory agrees well with empirical approaches to polarization detection. Their idea is to analyze the homogeneity and heterogeneity of social groups. On the one hand, the polarized group

must be highly homogeneous within itself but heterogeneous compared to other groups. In this approach, empirical tools focus on measuring the reliability of the differences between groups and the similarities of group members within themselves.
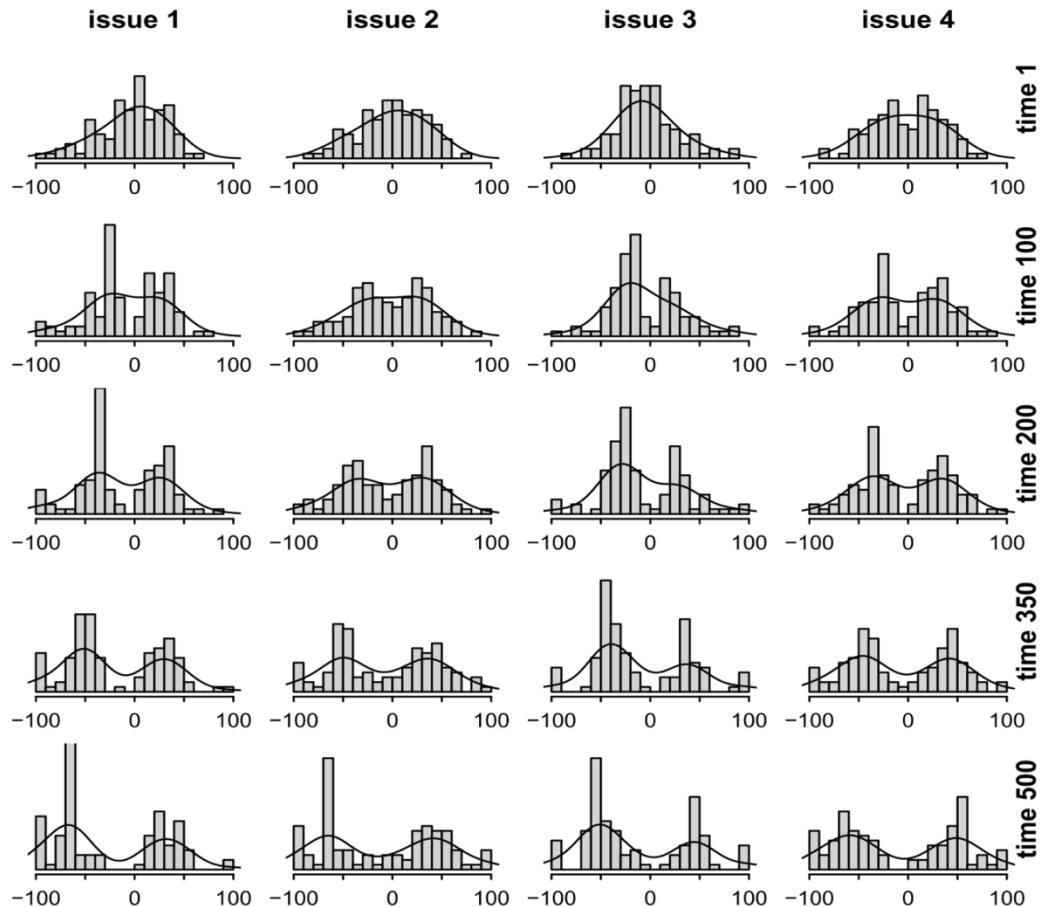

Fig. 2. *Baldassari-Bearman Polarization Simulations (Baldassarri, & Bearman, 2016)*

Also, a good mathematical approach to estimate polarization proposes econometric scientists Joan-Maria Esteban and Debraj Ray (Esteban, & Ray, 2010). They have constructed a social polarization index, but it is empirically tested only on income data. We can adapt it to social attitude data.

In the article Disambiguation of social polarization concepts and measures by Aaron Bramson, Patrick Grim, Daniel J. Singer, Steven Fisher, William Berger, Graham Sack & Carissa Flocken (Bramson et al., 2016) made a detailed grouping of approaches to a statistical assessment of group polarization. In total, they identified as many as ten approaches:

1. Spread: how wide is the attitudes field that is there in the system.

Dispersion: is about the whole distribution rather than only the extremes as is with spread.

Coverage: the number of attitudes that at least one person holds in the whole population.

Regionalization: However, the number of uncovered intervals (ignoring their size) is a distinct sense of variety and, hence, a different way to polarize distributions. The more empty regions of the attitude space, the greater the polarization is in the sense of regionalization.

Community fragmentation: to what extent the population can be divided into subpopulations. More groups mean greater polarization degree.

Role of groups in some senses of polarization: group identification can be a fragmentation measure.

Distinctness: how well we can distinguish different groups.

Group divergence: how distant the groups' attitudes are without accounting for their shapes.

Group consensus how groups are formed themselves.

Size parity: how the sizes of the groups affect our senses.

The advantage of the selected approaches is that they are well consistent with the above theoretical concepts. The only thing missing is the measurements by which to evaluate these divisions.

1.2. Lipset and Roccan's Theory of Cleavages in the Study of Social Polarization

The above concepts focus on the political dimension, but we take the position that political division is the result of hidden social divisions. Because of this, we consider it necessary to turn to concepts that try to explain the nature of such divisions. In this sense, the theory of Lipset and Rokkan is very interesting to us.

Cleavages theory of Lipset and Rokkan offers four dimensions of social conflict, which leads to the formation of social groups:

• center/periphery;

• state/church;

• workers/owners;

• land/industry.

These divisions were later supplemented by the winners/losers of globalization (Hanspeter Kriesi) and materialism/postmaterialism (Inglehart).

We can analyze these dimensions (all or part) on an empirical level through text analysis.

Advantages of using Lipset Rokkan's theory as a theoretical basis for studying social polarization:

– this theory offers specific dimensions around which social conflict and polarization develop;

– this theory is related to existing theories of polarization, which focus on party affiliations, as these dimensions of conflict are institutionalized, including through the formation of political parties.

Thus, it will allow us to solve several problems:

– the problem of connecting the concept of polarization with sociological theories: Lipset-Rokkan theory is fundamental and aims to explain the profound and long-term social rifts;

– the problem of the focus of empirical measurement: we get the dimensions of the analysis;

– the problem of the connection of previous studies of polarization (focused on policy analysis) with the sociological legacy.

Ukrainian researchers have successfully used Lipset-Roccane's theory to explain. The article «Group Nominations as an Instrument of Electoral Struggle: Ukrainian Discourse of 2019 Parliamentary Elections» uses the theory of social cleavages as a theoretical basis that explains the nature of differences between different parties and their supporters. The fundamental thesis of the Lipset-Roccan theory, on which researchers base it, is that party division is only a reflection of existing societal inequalities and differences. We agree with this thesis but want to expand it by connecting with the theories of social polarization (Boyko et al., 2021).

So, here we can propose our own scheme of social polarization. This scheme solves the problem of relations between social polarizations empirical studies and theory. In our mind, social polarization is the concept that can explain existing social distances and conflicts in society. We agree with Lipset-Rokkan that social distances (cleavages) have different dimensions. But for each dimension, we can calculate the index of polarization. It allows us to create a fully operationalized calculation system of social conflict estimation. The way that we have selected is the analysis of social media texts. It is a very accessible data source in which we can find all dimensions of conflict (center/periphery, state/church, workers/owners, land/industry). And these data are very suitable for automated text analysis techniques such as text mining.

CONCLUSION

The theory of social polarization grew out of the theory of social identification, which was used to explain why different parties' supporters differ so much from each other. The problem with these theories is that the first theory is socio-psychological, and the second is political science. There is currently the potential to use the concept of social polarization in sociological research. The Hobolt scheme, which includes three components: identification, differentiation, bias, is suitable to do this. We can quickly operationalize this scheme to use it in empirical studies, as demonstrated by the authors on Brexit analysis.

We already have experience in studying social phenomena through the latest methods of textual data analysis. The Baldassari-Bearman and Hobolt conceptions are suitable for use on text communication data sets. we are going to test them on future posts on Twitter, because we have an example of successful use of such data in sociological research (Kyrychenko, 2021).

Party affiliation is a factor that often serves as a basis for sustainable identification, but not always and not in every society. Hobolt's theoretical scheme enables, among other things, international comparative studies of social polarization.

As a result, we have identified four main challenges currently facing the sociological study of polarization:

1. Terminological Inconsistency. The authors often call the phenomenon under study social polarization, but there is the concept of «party polarization», «political polarization». The term «social polarization» is more commonly used now.

2. Political Science. Research focuses on the study of political processes, especially party confrontation in the United States. In most sources, the definition of social polarization is very narrow and boils down to the fact that it is merely a disagreement between Democrats and Republicans in the United States, which significantly limits the potential of this concept.

3. Americanization. The study of polarization often focuses on studying the American case, from which it can be wrongly concluded that it is only an American phenomenon.

4. Social Psychologism. The theoretical basis of modern theories of polarization is socio-psychological theories, particularly the theory of the social identity of Taifel, social comparison theory, persuasive arguments theory, self-categorization theory, and social decision scheme theory. The authors focus on the psychological aspects of polarization while largely ignoring the social factors that influence it.

These challenges can be overcome by turning to sociological theories of conflict and Lipset-Rokkan theory, which we propose as a theoretical basis for studying social polarization. We have adopted this scheme for social polarization estimation.

REFERENCES

Andersen, H. T. (2004). Spatial – Not Social Polarisation: Social Change and Segregation in Copenhagen. *Επιθεώρηση Κοινωνικών Ερευνών*, 113(113), 145. https://doi.org/10.12681/grsr.9356.

Baldassarri, D., & Bearman, P. (2016). *Dynamics of Political Polarization,* 72(5), 784–811.

Baum, S. (1997). Sydney, Australia: A global city? Testing the social polarisation thesis. *Urban Studies*, 34(11), 1881–1902. https://doi.org/10.1080/0042098975295.

Bliuc, A. M., McGarty, C., Reynolds, K., & Muntele, D. (2007). Opinion-based group membership as a predictor of commitment to political action. *European Journal of Social Psychology*, 37(1), 19–32. https://doi.org/10.1002/ejsp.334.

Boyko, D., Zaporozhchenko, R., Lytovchenko, A., Nekhaienko, O., & Yashkina, D. (2021). Group Nominations as an Instrument of Electoral Struggle: Ukrainian Discourse of 2019 Parliamentary Elections. *Sociological Studios*, 2(19), 41–52. https://doi.org/10.290 38/2306-3971-2021-02-41-52.

Bramson, A., Grim, P., Singer, D. J., Fisher, S., Berger, W., Sack, G., & Flocken, C. (2016). Disambiguation of social polarization concepts and measures. *Journal of Mathematical Sociology*, 40(2), 80–111. https://doi.org/10.1080/0022250X.2016.1147443.

Dalton, R. J. (2008). The Quantity and the Quality of Party Systems: Party System Polarization, Its Measurement, and Its Consequences. *Comparative Political Studies*, 41(7), 899–920. https://doi.org/ 10.1177/0010414008315860.

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization? *Public Opinion Quarterly*, 83(1), 114–122. https://doi.org/10.1093/poq/nfz003.

Esteban, B. Y. J., & Ray, D. (2010). On the Measurement of Polarization. *The Econometric Society Stable*, 62(4), 819–851. Retrieved April 12, 2020 from http://www.jstor.org/stable/2951734.

Hamnett, C. (1994). Social Polarisation in Global Cities: Theory and Evidence. *Urban Studies*, 31(3), 401–424. https://doi.org/10.1080/00420989420080401.

Hobolt, S., Leeper, T., & Tilley, J. (2020). Divided by the vote: affective polarization in the wake of the Brexit referendum. *British Journal of Political Science*, 51(4), 1476–1493. https://doi.org/10.1017/S0007123420000125.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science*, 22(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, 76(3), 405–431. https://doi.org/10.1093/poq/nfs038.

Johnston R. (2019). *Affective Polarization in the Canadian Party System 1988–2015*. Prepared for presentation at the 2019 annual meeting of the Canadian Political Science Association, Vancouver, BC, 4–6 June 2019. Retrieved May 10, 2020 from https://rgcjohnston.files.wordpress.com/2019/08/johnston-polarization-cpsa-2019-text.pdf.

Kyrychenko, R. (2021). Typology of Tasks of Machine Analysis of Texts in Contemporary Sociology. *Sociological Studios*, 2(19), 53–62. https://doi.org/10.29038/2306-3971-2021-02-41-48.

Lauka, A., McCoy, J., & Firat, R. B. (2018). Mass Partisan Polarization: Measuring a Relational Concept. *American Behavioral Scientist*, 62(1), 107–126. https://doi.org/10.1177/0002764218759581.

Levendusky, M. S., & Malhotra, N. (2013). The Effect of False Polarization: Are Perceptions of Political Polarization Self-Fulfilling Prophecies? *Annual Meeting and Exihibition of the American Political Science Association*, 2(6), 1–34.

Luttig, M. D. (2017). Authoritarianism and affective polarization: A new view on the origins of partisan extremism. *Public Opinion Quarterly*, 81(4), 866–895. https://doi.org/10.1093/poq/nfx023.

Mason, L. (2015). "I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science*, 59(1), 128–145. https://doi.org/10.1111/ajps.12089.

Mason, L. (2018). Ideologues without issues: The polarizing consequences of ideological identities. *Public Opinion Quarterly*, 82(S1), 280–301. https://doi.org/10.1093/poq/nfy005.

McGarty, C., Bliuc, A. M., Thomas, E. F., & Bongiorno, R. (2009). Collective action as the material expression of opinion-based group membership. *Journal of Social Issues*, 65(4), 839–857. https://doi.org/10.1111/j.1540-4560.2009.01627.x.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178. https://doi.org/10.1002/ejsp.2420010202.

Tworzecki, H. (2019). Poland: A Case of Top-Down Polarization. *Annals of the American Academy of Political and Social Science*, 681(1), 97–119. https://doi.org/10.1177/0002716218809322.