Hybrid CoE Paper 24

Social identities and democratic vulnerabilities: Learning from examples of targeted disinformation

Hybrid CoE

Elsa Hedling – April 2025

Hybrid CoE Papers are finalized pieces of analysis on a topic related to hybrid threats, based on one or several research questions. They may be either conceptual analyses or based on a concrete case study with empirical data.

The European Centre of Excellence for Countering Hybrid Threats
tel. +358 400 253800 www.hybridcoe.fi

ISBN 978–952–7591–22–2 (web)
ISBN 978–952–7591–23–9 (print)
ISSN 2670–2053 (web)
ISSN 2814–7227 (print)

April 2025

Cover photo: Juergen Faelchle / Shutterstock.com

Hybrid CoE's mission is to strengthen its Participating States' security by providing expertise and training for countering hybrid threats, and by enhancing EU-NATO cooperation in this respect. The Centre is an autonomous hub for practitioners and experts, located in Helsinki, Finland.

Contents

Summary

Hybrid threat actors are known to use tactics such as misinformation and disinformation to deliberately fuel mistrust and polarization in order to destabilize democracies. They manipulate essential democratic processes, such as equal opportunities for political participation, by turning them into security vulnerabilities. The exploitation of identity politics, enabled and intensified by the modern information environment, has become a way of weaponizing inclusive democracy. Hybrid threats emerge and evolve rapidly and can reproduce intergroup conflict along multiple and opposing politicized lines. False narratives, especially when emotionally charged, can quickly spread in ways that challenge authorities by exposing deeply ingrained grievances and perceived inequalities and injustices that erode societal trust among targeted communities. This paper outlines the intersection of social identities and hybrid threats by defining key terminology, exploring tactics, and highlighting case studies that provide examples of disinformation targeting social identities. The aim is to deepen understanding of vulnerabilities and provide actionable insights to help policymakers and practitioners manage and deter these threats in an effort to bolster democratic resilience.

1. Introduction

Democracies are characterized by pluralism and multiple processes of deliberation and negotiation among different groups that represent various socio-political perspectives. It is this source of pluralism that ensures diverse political options and meaningful alternatives to policies that are available for public debate and collective decision-making. The strength of liberal democracies lies in their ability to sustain these processes, enabling societies to resolve conflicts between different groups without resorting to violence or structural forms of oppression. The representation of multiple social identities is therefore an important sign of the vitality of democracies. It is the ability to foster meaningful dialogue and reach compromises that makes democracies resilient to political turbulence by allowing them to adapt to challenges. However, the same openness and access to rights such as freedom of speech and a diverse media landscape can be exploited by malicious actors that seek to undermine public trust, spread misinformation and disinformation, and sow the seeds of hostility in societies. Polarization can further drive division and the construction of an 'us' versus 'them' mentality, leading to heightened perceptions of conflict. When pluralism is manipulated to amplify polarization, conflicts can quickly escalate into crises that threaten the stability and status of democratic systems.[1] Contestation over political interests that mobilize social identities in different ways can therefore be exploited for the purpose of dividing societies. Hybrid threat actors, including state and non-state entities known to combine conventional and unconventional tactics such as disinformation campaigns, cyberattacks and economic coercion,[2] increasingly exploit this vulnerability. Social identities in democracies can therefore become sources of both vulnerability and resilience in the context of hybrid threats.

This paper focuses on the role of social identities in targeted disinformation efforts. When threat actors exploit divisions around political issues and interests, social identities become both drivers and targets of polarization. By concentrating on this particular aspect of democratic vulnerability, this paper aims to inform policymakers, officials and practitioners about the conditions and risks under which contestation and polarization can be exploited and amplified, with a view to better addressing and mitigating these challenges. The paper begins with a brief overview of key terminology, which serves as a framework for the subsequent illustrative analysis of known cases of targeted disinformation. Specifically, it introduces and defines the role of social identities in democratic politics, illustrating how they can contribute to processes of cooperation as well as serve as drivers of conflict in democracies. The paper then considers the conditions under which threat actors may exploit domestic sources of contestation to weaken democratic institutions, sow hostility between groups or undermine societal trust. It provides examples of domestic contestation or polarization in political processes or crises where threat actors have been found to be involved in the diffusion of

1 Jennifer McCoy, Tahmina Rahman, and Murat Somer, 'Polarization and the Global Crisis of Democracy: Common Patterns, Dynamics, and Pernicious Consequences for Democratic Polities', *American Behavioral Scientist* 62, no. 1 (2018): 16–42.

2 Hybrid CoE, 'Hybrid Threats as a Concept', 2024, https://www.hybridcoe.fi/hybrid-threats-as-a-phenomenon/.

disinformation targeting social identities. These illustrative cases are neither exhaustive nor indicative of the distribution of social identities targeted, but rather chosen to demonstrate the various ways in which social identities (in multiple intersectional combinations) can be exploited as vulnerabilities to undermine social cohesion in democracies. The paper concludes by stressing the value of informed analysis of democratic vulnerabilities and the conditions under which they may be exploited, as well as the importance of proactive efforts to strengthen democratic resilience against hybrid threats.

2. Key terminology

In democratic systems, **social identities** refer to the ways in which individuals and groups in a given society define themselves and are defined by others within that same society. These processes of identification are based on affiliations and attributes that are recognized as socially meaningful in the political system. Social identities are related to and partly overlap with political identities, which are more closely aligned with ideology in the sense that they are based on social characteristics rather than political alignment alone.[3] This implies that social identities are group identities that exist outside of political systems. Social identities can still play an important role in fostering political identities, and political ideology can be one of many social identities in a democratic system.

The affiliations and attributes that shape social identities include religion, ethnicity, nationality, gender, socioeconomic status and more. Social identities may influence political behaviour, attitudes towards policy and patterns of interaction within democratic societies. They can also be the source of political mobilization through activism or civil society engagement.[4] Groups with shared social identities often mobilize to advocate for shared interests and policy preferences. As such, social identities are important sources of cooperation in democratic systems. For instance, despite widespread secularization, religious identity remains important to party affiliations in many democratic systems.[5] At the same time, religion is also a known source of bias and discrimination, often contributing to tensions and conflict between groups. Hence, while social identities are important sources of cooperation within groups, they can also contribute to conflict and to the perception of other groups as threatening to one's own. Social identity theory explains this through the social tendency of individuals to categorize themselves according to '**in-groups**' (groups to which they belong) and '**out-groups**' (groups to which they do not belong).[6] These processes of **social categorization** have been found to lead to favouritism towards one's own in-group and a negative bias towards the out-group.[7] Biased behaviour, which is often unconscious, entails categorizing other people according to characteristics shared within a group (e.g., ethnicity, religion or gender) in ways that simplify social perception. The effect thereof can be an increased emphasis on differences between groups (intergroup differences) and less emphasis on differences within groups (intragroup similarity).[8]

The following list defines social identities that are known to be influential in processes that can lead to both cooperation and conflict in democratic politics:

3 Leonie Huddy, 'From Social to Political Identity: A Critical Examination of Social Identity Theory', *Political Psychology* 22, no. 1 (2001): 127–156.

4 David E Campbell, 'Social Networks and Political Participation', *Annual Review of Political Science* 16, no. 1 (2013): 33–48.

5 Oddbjørn Knutsen, 'Religious Denomination and Party Choice in Western Europe: A Comparative Longitudinal Study from Eight Countries, 1970–97', *International Political Science Review* 25, no. 1 (2004): 97–128.

6 Henri Tajfel, 'An Integrative Theory of Intergroup Conflict', in *The Social Psychology of Intergroup Relations* (Brooks/Cole, 1979).

7 John C Turner, Rupert J Brown, and Henri Tajfel, 'Social Comparison and Group Interest in Ingroup Favouritism', *European Journal of Social Psychology* 9, no. 2 (1979): 187–204.

8 Henri Tajfel, *Social Identity and Intergroup Relations*, vol. 7 (Cambridge University Press, 2010).

- **Ethnic, cultural and religious identities** relate to an individual's sense of belonging to a distinct ethnic group, or cultural or religious community. These identities, which may be separate or overlapping, can influence political affiliations and policy preferences, including how in-group members perceive issues such as immigration, minority rights and national identity.[9] Religious identity has also been known to influence political preferences for policies related to morality, education, sexual and reproductive health and rights, among others.
- **Gender identity and sexual orientation** includes an individual's perception of their own gender, which may influence attitudes towards gender equality, reproductive rights, and labour rights. It can also influence perceptions of political representation.[10] Sexual orientation, which refers to an individual's romantic or sexual attraction, can also cause social stigma, experiences of which may shape voting behaviour.[11]
- **Socioeconomic status** refers to the combination of professional identity, educational background, access to economic resources (networks) and social status in relation to others in a given society. Socioeconomic status is known to influence political attitudes and policy preferences regarding issues such as taxation, welfare, environmentalism and other areas of economic policy.[12]
- **Political ideology** may feature as a core aspect of an individual's social identity by taking precedence in shaping attitudes towards the role of the government, individual versus group rights, and aspects of social justice. Political ideology is often the source of party affiliation and voting behaviour. It can also influence perceptions of broader social issues such as economic policy, civil rights and international politics.

There are many other social identities or combinations thereof that contribute to both pluralism and polarization in democratic systems. The fact that social identities often overlap and intersect is an important aspect of the complexity of social belonging in politics. The term **intersectional identities** is used to highlight this complexity in how attributes such as race, gender and social class can overlap to influence an individual's social experiences, and how they may shape political behaviour. It is also important to note that the ways in which social identities influence political behaviour are not static over time. Research on voting preferences and social identities has produced mixed results, which are often attributed to varying degrees of political polarization at different times across political contexts.[13]

9 Jan E Leighley, *Strength in Numbers?: The Political Mobilization of Racial and Ethnic Minorities* (Princeton University Press, 2001).

10 Simon Bornschier et al., 'How "Us" and "Them" Relates to Voting Behavior—Social Structure, Social Identities, and Electoral Choice', *Comparative Political Studies* 54, no. 12 (2021): 2087–2122.

11 Michael Hunklinger and Philipp Kleer, 'Why Do LGB Vote Left? Insight into Left-Wing Voting of Lesbian, Gay and Bisexual Citizens in Austria', *Electoral Studies* 87 (2024): 102727.

12 Jazmin L Brown-Iannuzzi, Kristjen B Lundberg, and Stephanie McKee, 'The Politics of Socioeconomic Status: How Socioeconomic Status May Influence Political Attitudes and Engagement', *Current Opinion in Psychology* 18 (2017): 11–14.

13 Ryan Strickler, 'Deliberate with the Enemy? Polarization, Social Identity, and Attitudes toward Disagreement', *Political Research Quarterly* 71, no. 1 (2018): 3–18.

In this context, the term **political polarization** refers to an increasing division between groups (in terms of an ideological divide or fragmentation in society) caused by in-group perspectives and processes of social categorization.[14] Furthermore, identity-driven feelings of political affiliation are known as **affective polarization**, a process whereby in-group members harbour positive feelings towards their own group and negative feelings towards the out-group, which may be a political party or an otherwise defined group with opposing interests.[15] Social identity theory thus explains phenomena such as pluralism, political contestation and polarization by paying attention to how an individual's perception of social categories shapes both political and emotional bonds within groups and between groups in society. These bonds and perceptions of intergroup threat intensify in moments of uncertainty, such as a political crisis. Increased inter-group threat perceptions also lead to greater social distances within a political system.[16] This means that these processes can weaken **social cohesion** in terms of the degree of unity, solidarity and mutual understanding of diversity as a source of strength in a political community.[17]

14 Delia Baldassarri and Peter Bearman, 'Dynamics of Political Polarization', *American Sociological Review* 72, no. 5 (2007): 784–811.

15 Turner, Brown, and Tajfel, 'Social Comparison and Group Interest in Ingroup Favouritism'.

16 Emma A Renström, Hanna Bäck, and Royce Carroll, 'Threats, Emotions, and Affective Polarization', *Political Psychology* 44, no. 6 (2023): 1337–1366.

17 David Schiefer and Jolanda Van der Noll, 'The Essentials of Social Cohesion: A Literature Review', *Social Indicators Research* 132 (2017): 579–603.

3. Hybrid threats and social identities

Hybrid threats refer to actions conducted by either state or non-state actors with the goal of undermining or harming a target using overt and covert military and non-military means. Such threats aim to undermine the target, which may be a state, a government, or an institution, through multiple means that are often combined, including, but not limited to, foreign information manipulation, interference, cyberattacks, economic influence or coercion, covert operations, coercive diplomacy or threats to use military force. Hybrid threats employ a wide range of known tactics, techniques and procedures with different goals.[18] In contrast to traditional warfare, hybrid threats exploit civil (rather than military) vulnerabilities in target societies, such as social divisions, weak infrastructure, or political instability, to increase confusion or sow distrust, impacting the ability to respond effectively to the crisis at hand. This subtle and often covert nature of hybrid threats makes them difficult to assess, attribute and counter, as they often pass just under the radar of aggression.[19] Hybrid threats effectively blur the lines between interstate wars, the escalation of interstate conflicts and domestic conflicts. The role of social identities as sources of pluralism and cooperation, and as triggers of tension and conflict, makes them particularly vulnerable to exploitation by hybrid threat actors that seek to undermine and manipulate essential democratic processes. These hybrid threat actors, most notably Russia, China and Iran, are known to exploit domestic sources of division by spreading or amplifying misinformation, disinformation and divisive narratives for these purposes. By manipulating democratic processes, they seek to increase social fragmentation to weaken the capacity of the targeted state actors. The manipulation is seldom ideologically aligned. For instance, a common strategy is to amplify opposing sides in political debates simultaneously, as has been the case in the context of the Black Lives Matter movement in the United States, and debates around LGBT+ rights and migration.[20] The goal of this strategy is therefore not to sway political opinion but to increase division, undermine trust in the government and other institutions, such as the police, and to create a conflictual environment in which dialogue and compromise become increasingly difficult to achieve.[21]

Political polarization around issues pertaining to social identity tends to increase in times of crisis or widespread hardship. For example, economic downturns, including financial crises, inflation and the implementation of austerity measures, can create fertile ground for

18 Hybrid CoE, 'Hybrid Threats as a Concept'.

19 Elsa Hedling and Hedvig Ördén, 'Disinformation, Deterrence and the Politics of Attribution', *International Affairs*, forthcoming 2025.

20 Deen Freelon et al., 'Black Trolls Matter: Racial and Ideological Asymmetries in Social Media Disinformation', *Social Science Computer Review* 40, no. 3 (2022): 560–578; Gunhild Hoogensen Gjørv and Outi Jalonen, 'Identity as a Tool for Disinformation: Exploiting Social Divisions in Modern Societies', Hybrid CoE Strategic Analysis 34 (Hybrid CoE, November 2023), https://www.hybridcoe.fi/wp-content/uploads/2023/11/20231108-Hybrid-CoE-SA-34-Identity-as-a-tool-for-disinformation-WEB.pdf.

21 Jean-Baptiste Jeangène Vilmer, 'The "Macron Leaks" Operation: A Post-Mortem' (Atlantic Council, 2019), https://www.atlanticcouncil.org/wp-content/uploads/2019/06/The_Macron_Leaks_Operation-A_Post-Mortem.pdf.

increased polarization.[22] The distribution of economic burdens and the perceived lack of contribution from certain groups may fuel pre-existing grievances along identity lines. Regardless of whether these perceptions are justified, such experiences tend to heighten tensions between groups, leading to competition for attention and resources. These dynamics can further drive social divisions and suspicions of unfair treatment. For example, the arrival of refugees in need of state assistance can cause resentment among groups already receiving insufficient support from their government. These communities may feel that their interests and needs are being neglected in favour of those of other groups. Such feelings can lead to frustration and further deepen intergroup threat perceptions, as groups see themselves as rivals competing for limited governmental support.

These dynamics, characterized by the same pluralism and diversity as processes of deliberation and cooperation in democracies, can become vulnerabilities that threat actors exploit. For instance, frustration over the unfair distribution of resources can lead to the scapegoating of certain groups, providing opportunities to weaponize division. Tactics exploiting group grievances are not only divisive, but often more effective in times of crisis and uncertainty, when individuals gravitate towards political simplification and emotional narratives that align with their in-group identity.[23] Research has shown that in situations of uncertainty, individuals are more prone to align their perceptions and actions with the dominant in-group beliefs.[24] When faced with ambiguity, individuals adopt shared interpretations and narratives about "problems" and "solutions" in search of cognitive clarity. These cognitive shortcuts may be biased and lead to support for discriminatory statements or practices based on misleading information that vilifies out-group members.[25] Uncertainty and ambiguity further increase the need for group validation, which may lead individuals in both the dominant in-group and the targeted out-group to seek consensus by endorsing statements about other groups that may be both prejudiced and incorrect.[26] These

22 Mark Blyth, *Austerity: The History of a Dangerous Idea* (Oxford University Press, 2013); Evelyne Hübscher, Thomas Sattler, and Markus Wagner, 'Does Austerity Cause Polarization?', *British Journal of Political Science* 53, no. 4 (2023): 1170–1188.

23 Tobias Widmann, 'Fear, Hope, and COVID-19: Emotional Elite Rhetoric and Its Impact on the Public during the First Wave of the COVID-19 Pandemic', *Political Psychology* 43, no. 5 (2022): 827–850; Thomas Kessler and Amélie Mummendey, 'Is There Any Scapegoat around? Determinants of Intergroup Conflicts at Different Categorization Levels', *Journal of Personality and Social Psychology* 81, no. 6 (2001): 1090; Micheal O'Flynn, Lee F Monaghan, and Martin J Power, 'Scapegoating during a Time of Crisis: A Critique of Post-Celtic Tiger Ireland', *Sociology* 48, no. 5 (2014): 921–937.

24 Joanne R Smith et al., 'Uncertainty and the Influence of Group Norms in the Attitude–Behaviour Relationship', *British Journal of Social Psychology* 46, no. 4 (2007): 769–792.

25 Judit Szakacs and Eva Bognar, 'The Impact of Disinformation Campaigns about Migrants and Minority Groups in the EU', (Policy Department for External Relations, Directorate General for External Policies of the Union, 2021), https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/INGE/DV/2021/07-12/IDADisinformation_migrant_minorities_EN.pdf.

26 Tajfel, *Social Identity and Intergroup Relations*.

situations of uncertainty create the perfect conditions for threat actors to exploit existing divisive narratives.[27]

Increased migration flows can heighten political polarization and intergroup threat perception beyond frustration over competition for resources. In addition to introducing new interests and needs into a society, migration is associated with cultural, social and demographic changes that may increase existing tensions between groups in society. These changes may impact debates on issues such as national identity, integration and security by challenging core societal values such as religious freedom or gender equality. Major migration flows may therefore also trigger identity-based fears among host populations when cultural values, job security or national identity are perceived to be threatened.[28] Domestic actors have also been known to exploit these concerns by framing migrants or minority groups as a cultural or economic threat.[29] These domestic political narratives may also further drive divisions between groups and shift the focus of political discourse from policy alternatives to more emotionally charged conflicts over identity and belonging. Moreover, a public health crisis such as the Covid-19 pandemic can deepen political polarization by further exposing inequalities and prejudices between groups, thereby catalyzing debates about government legitimacy, and the status of expertise and science.[30] When social inequalities become contested, different groups may perceive government responses differently based on their in-group identity. Another category of events that increases the risk of vulnerability concerns political turbulence, such as government crises or post-election controversies, which have become more common in Western Europe and the United States.[31] Similar to other conflicts that vilify the out-group in various ways, such events, especially if they occur frequently, may erode trust both between groups and in government institutions. These processes of social categorization, combined with uncertainty and low trust in institutions, create ample opportunities for threat actors to exploit the situation in order to distort and destabilize the targeted societies. Hence, when identity-based grievances and social divisions between groups are already evident, further turbulence – through crisis or uncertainty or both – creates ideal conditions for malicious manipulation.

27 Madhavi Reddi, Rachel Kuo, and Daniel Kreiss, 'Identity Propaganda: Racial Narratives and Disinformation', *New Media & Society* 25, no. 8 (2023): 2201–2218.

28 Jens Hainmueller and Daniel J Hopkins, 'Public Attitudes toward Immigration', *Annual Review of Political Science* 17, no. 1 (2014): 225–249.

29 Cas Mudde, *Populist Radical Right Parties in Europe*, vol. 2007 (Cambridge University Press, 2007).

30 Sophie L Vériter, Corneliu Bjola, and Joachim A Koops, 'Tackling COVID-19 Disinformation: Internal and External Challenges for the European Union', *The Hague Journal of Diplomacy* 15, no. 4 (2020): 569–582.

31 Bernard L Fraga, Zachary Peskowitz, and James Szewczyk, 'Can Elite Allegations of Election Fraud Demobilize Supporters?', *Political Behavior*, 2024, 1–25.

4. Learning from examples: 2016–2024

The following four case studies have been chosen to illustrate how social identities and intergroup threat perceptions have been mobilized or exploited by hybrid threat actors with an interest in destabilizing democracies. These cases differ in context and scope, serving to exemplify the key terminology and relationship between social identities and hybrid threats in real-life scenarios. There are many other examples of the ways in which social identities are targeted, or tensions between groups exploited in different democratic contexts and involving different threat actors. The selected examples serve to show the variation and highlight the complexity of intersectional identities when group interests are co-opted and politicized for different agendas.

4.1 Refugee crises in Germany: Ethnic, cultural, and religious identity and political ideology

One of the most widely known and discussed examples of a European disinformation campaign is the "Lisa case", which drew attention to the problem in Germany and ultimately shaped Germany's response to Russian foreign interference. The Lisa case also demonstrates how social identities were mobilized in a German context characterized by an ongoing refugee crisis (involving an influx of people with perceived ethnic and religious identities), a Russian minority group (cultural identity), and a government that was at the time reluctant to alienate Russia (political ideology).

It was in January 2016 that news media first reported that a 13-year-old German girl of Russian origin, "Lisa", claimed to have been abducted and raped by a group of migrant refugees in Berlin. The reports first came to light on Russian television channel 1, after which Russian foreign media circulated the reports further. Lisa was allegedly kidnapped on her way to school and gang-raped by "southern-looking" asylum seekers. The story was debunked rather quickly by the German police (who found that she had in fact been with a friend that night and found no medical evidence of rape).[32] However, the story had already gone viral on social media after being spread by domestic right-wing groups as well as new Russian state-controlled media outlets. The false narrative had caused considerable outrage and heightened anti-refugee sentiments at a time of crisis (Germany and the EU were grappling with the large influx of refugees from the Middle East).[33]

The circulation of Lisa's story sparked a large anti-immigration protest march, which was reported by Russian and German news media, resulting in public statements by Russian Foreign Minister Sergei Lavrov condemning Germany's handling of the refugee crisis. Analyses have found that the story was deliberately fabricated to align with the Kremlin's narrative at the time of questioning Europe's ability to cope with the influx of refugees and portraying the EU as "soft and weak" as opposed to Russia as strong and stable.[34]

32 Christina La Cour, 'Theorising Digital Disinformation in International Relations', *International Politics* 57, no. 4 (2020): 704–723.

33 Stefan Meister, 'The "Lisa Case": Germany as a Target of Russian Disinformation', *Nato Review* 25 (2016): 2016.

34 Jakub Janda, 'The Lisa Case STRATCOM Lessons for European States', Security Policy Working Paper, No. 11/2016 (Federal Academy for Security Policy, 2016), https://www.baks.bund.de/sites/baks010/files/working_paper_2016_11.pdf.

The Lisa case was particularly significant because of the timing of the circulation of the story in the context of heightened political tensions following Chancellor Angela Merkel's controversial position during the refugee crisis. Merkel faced mounting criticism for Germany's open-door policy, with the tactics surrounding the Lisa story exploiting not only the divisions between Germans and refugees but also other fears in German society. By playing on fears of unregulated migration and Muslim immigration in particular, the narratives that surrounded this case served to intensify the intergroup threat perception by capitalizing on anxieties about multiculturalism and religious rights in Germany. The disinformation campaign around the Lisa case triggered emotions in line with affective polarization by exploiting public emotions and fears during a time charged with political contestation and uncertainty. Although the alleged story was debunked and links to Russia were proven, this case remains a stark warning of how hybrid threats, and in this case disinformation, can manipulate public perceptions and influence political outcomes when social factions present a vulnerability.

What made the Lisa story particularly potent in its circulation on online platforms was the fact that it was reported shortly after a series of news reports about attacks on women by migrant men during New Year's Eve in Cologne.[35] The Lisa story was thus also a tactic to further fuel emotions and anti-migrant sentiments at their peak by mixing a false and planted story with reactions to actual events. This case therefore also highlights the importance of strategic timing in times of political turbulence, which in this case enabled the exploitation of collective fears linked to both in-group identity and out-group threat construction.

4.2 Anti-establishment sentiments in France: Socioeconomic status, class and political ideology

The Gilets Jaunes (Yellow Vests) protests in France in 2018 and 2019 became the target of interference and disinformation campaigns orchestrated by both domestic and foreign threat actors that sought to exploit the state of social unrest to destabilize France. The social identities targeted in this case were initially framed in terms of socioeconomic status and class, but evolved into a rural-urban divide and broader anti-establishment sentiments. The protests, which began in opposition to a fuel tax increase quickly escalated into riots and clashes over economic inequality, populism, mistrust in government and more, causing deep divisions within the French population.[36]

The Gilets Jaunes protests, named after the high-visibility vests worn by the initial protesters, began as a loosely organized protest movement in France in October 2018. The protests got underway on social media following the announcement of an increase in the country's carbon tax (TICPE), and by November, numerous protests were taking place by means of roadblocks, occupations of roundabouts, and

35 Jeffrey Mankoff, 'Russian Influence Operations in Germany and Their Effect' (Center for Strategic and International Studies, 2020), https://www.csis.org/analysis/russian-influence-operations-germany-and-their-effect.

36 Élie Guckert, 'Manufacturing Dissent: RT France's Challenge in a Brand-New Media Landscape' (Centre for Democratic Integrity, 2022), https://democratic-integrity.eu/manufacturing-dissent-rt-frances-challenge-in-a-brand-new-media-landscape-by-elie-guckert/.

demonstrations. The movement first drew support from residents in rural and suburban areas, who were particularly affected by the tax increase on top of the high cost of living due to austerity measures, but it quickly spread to major cities. The rapid spread of the movement and the heightened sense of urgency as a "people's revolt" led to violent clashes with the police, particularly visible on the Avenue des Champs-Élysées in Paris.[37]

With new protests every week, the Gilets Jaunes voiced their demands for political reform and greater social justice. The violent demonstrations heightened frustration with the government's lack of response, leading to growing hostility centred on President Emmanuel Macron. Media coverage focused on the gap between the economic burdens and stresses of the protesters and the disinterested political elites, framing the movement as a working-class uprising. Critical coverage of the French government and Macron continued to spin the story of a government indifferent to the struggles of the working class. The growing tensions and unrest caused by the protests and media narratives were further amplified by a surge in social media posts and false news stories that fuelled anti-elite sentiments. For example, false stories that gained traction on social media included unrelated photographs of Macron socializing with Middle Eastern elites while people faced violent police suppression. Some of the images depicting political violence were also proven false, as they were found to show other, unrelated protests that had taken place in Spain several years prior to the Gilets Jaunes protests.[38]

The French government rolled back on the carbon tax increase and raised the minimum wage in France, but tensions were already high as protesters continued to push for reform. The situation worsened when the French government attempted to increase crowd control measures, leading to controversy and criticism in the French parliament.[39] This state of unrest created a "perfect storm" for Russian-affiliated media outlets such as RT and its French branch, RT France, as well as Sputnik. These media outlets specifically served to link the Gilets Jaunes protesters with other anti-elite movements active in France in a collective revolt against "oppressive and authoritarian policies".[40] This convergence tactic echoed the strategies employed by the same Russian media outlets during the "Macron Leaks", the hacked and leaked emails from Macron's 2017 presidential campaign. The Macron Leaks were an interference campaign that combined hacked and

37 Doron Shultziner and Irit S Kornblit, 'French Yellow Vests (Gilets Jaunes): Similarities and Differences with Occupy Movements', vol. 35 (Sociological Forum, Wiley Online Library, 2020), 535–542.

38 Avaaz Reports, 'Yellow Vests Flooded by Fake News', March 2019, https://avaazimages.avaaz.org/Report%20Yellow%20Vests%20FINAL.pdf.

39 Jean-Baptiste Racquin and Manon Rescan, 'Loi Anticasseurs : L'article Le plus Critiqué Censuré Par Le Conseil Constitutionnel', *Le Monde* 2019 (April 2019), https://www.lemonde.fr/politique/article/2019/04/04/l-article-phare-de-la-loi-anticasseurs-censure-par-le-conseil-constitutionnel_5445806_823448.html.

40 Tobias Lemke and Michael W Habegger, 'Foreign Interference and Social Media Networks: A Relational Approach to Studying Contemporary Russian Disinformation', *Journal of Global Security Studies* 7, no. 2 (2022): ogac004.

leaked emails with disinformation by mixing in fake emails.[41] While the Macron Leaks were deemed largely unsuccessful in influencing the election results, the disinformation campaign demonstrated how international actors could interact with French alternative media actors mobilized to disseminate disinformation, including fake photographs and videos, to susceptible domestic audiences.[42] The Alliance for Securing Democracy (ASD) identified up to 600 different accounts with links to the Kremlin alongside media outlets such as RT and Sputnik, which were involved in efforts to amplify divisive narratives during the Gilets Jaunes protests, often using fake photographs and promoting the idea of an ongoing systemic collapse.[43] A report published by AVAAZ showed that RT France was the most viewed YouTube channel for Gilets Jaunes-related stories at the height of the crisis, surpassing the collective views of French news media outlets *Le Monde*, *L'Obs*, *Le Huffington Post*, *Le Figaro* and *France 24*.[44]

At the peak of the Gilets Jaunes protests, images and reports of police violence and brutality stoked further outrage. While violent incidents, many reported injuries and some verified cases of police brutality did occur, with some officers even convicted of these crimes,[45] fake accounts and manipulated stories and photographs exaggerated the violence. This served to distort the scale of the clashes between protesters and the police. Russian state media used the situation to portray Macron's government as overly oppressive and out of touch with the French people. These narratives both perpetuated the state of unrest and further undermined trust in democratic institutions.[46]

In the case of the Gilets Jaunes protests, polarization between multiple social groups and a sustained crisis narrative were exploited by threat actors to deepen divisions. In this case, the conflict began with economic grievances and the unfair burden on the rural and working classes, but was relatively quickly reframed as a struggle between the "real" French people and the out-of-touch urban elite. This narrative, which aligned with the populist rhetoric and anti-establishment sentiments among alternative political factions, transformed the Gilets Jaunes move-

41 Jeangène Vilmer et al., 'Lessons from the Macron Leaks', *Hacks, Leaks and Disruptions: Russian Cyber Strategies*, 2018, 75–84.

42 Jean-Baptiste Jeangène Vilmer, 'The "Macron Leaks" Operation: A Post-Mortem' (Atlantic Council, 2019), https://www.atlanticcouncil.org/wp-content/uploads/2019/06/The_Macron_Leaks_Operation-A_Post-Mortem.pdf.

43 Alliance for Securing Democracy, 'Russian State Media, Including RT and Sputnik, Amplify and Heavily Promote the Yellow Vest Protest Movement in France on Social Media', 2019 2018, https://securingdemocracy.gmfus.org/incident/russian-state-media-including-rt-and-sputnik-amplify-and-heavily-promote-the-yellow-vest-protest-movement-in-france-on-social-media/.

44 Avaaz Reports, 'Yellow Vests Flooded by Fake News'.

45 'Deux Policiers Mis En Examen Pour Violences Envers Des « gilets Jaunes », Dont Jérôme Rodrigues, Qui a Perdu Son Œil', *Le Monde*, February 2021, https://www.lemonde.fr/societe/article/2021/02/10/deux-policiers-mis-en-examen-pour-violences-envers-des-gilets-jaunes-dont-jerome-rodrigues_6069500_3224.html.

46 Maxime Audinet, Colin Gérard, and Thomas Delage, 'Quelle Influence Russe En France?', *Diplomatie*, no. 116 (2022): 82–86.

ment into a broader symbol of resistance against economic policies, globalism, immigration, and other cultural issues that were absent from the protest's original purpose and demands. Hybrid threat actors actively used this identity fault line to further polarize French politics. While France is known for its tradition of protest movements, this particular protest provided suitable conditions to fuel intergroup threat perceptions through an "us versus them" mentality and dissent from both far-right and far-left groups, who saw the protests as opportunities to advance their anti-establishment agendas. This deliberate manipulation of identity politics further reinforced the framing of the protesters as defenders of national identity against an elite accused of betraying France through policies in favour of globalization and multiculturalism.[47] This framing meant that even when the protesters' initial demands were met, the riots continued. These conditions thus enabled threat actors to exploit what started out as a democratic protest by radicalizing certain elements within the movement, demonstrating how the intersection of economic grievances, social divisions and violent clashes can expose democracies to interference.

4.3 Attacks on women and sexual minorities in visible roles: Gender identity and sexual orientation

Recently published analyses of cases of foreign interference during democratic elections have reported how prominent female political candidates in US presidential elections, such as Hillary Clinton (in 2016) and Kamala Harris (in 2020 and 2024), were disproportionately targeted by disinformation explicitly based on their gender identity. These disinformation narratives exploit biased attitudes towards women.[48] This pattern of targeted disinformation thus amplifies or exploits polarization around gender norms, including issues such as gender equality, sexual and reproductive health and rights, and LGBT+ rights. In recent years, it has been largely associated with Russia's tactics of interference, whereby the targeting of gender identities aligns with the doctrine of traditional values. However, instances of this pattern in disinformation campaigns have been reported worldwide.[49]

By diffusing or amplifying narratives that disqualify women and gender non-conforming individuals from public roles, individuals who share these social characteristics may be discouraged from seeking political office, leadership roles, or other public positions (such as journalists, civil society leaders,

47 Lucia Posteraro, Timothy Peace, and Marius Nyquist Pedersen, 'Riding the Yellow Wave: The Online Populist Communication of Rassemblement National (RN) Leaders in Response to the Gilets Jaunes Protests and the 2019 European Elections', *Information, Communication & Society* 27, no. 8 (2024): 1712–1735.

48 Lucina Di Meco and Kristina Wilfore, 'Gendered Disinformation Is a National Security Problem' (Brookings Institution 2021), https://www.brookings.edu/articles/gendered-disinformation-is-a-national-security-problem/; Samantha Bradshaw and Amélie Henle, 'The Gender Dimensions of Foreign Influence Operations', *International Journal of Communication* 15 (2021): 23.

49 United Nations, 'Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, Irene Khan', August 2023, https://documents-dds-ny.un.org/doc/UNDOC/GEN/N23/233/65/PDF/N2323365.pdf?OpenElement.

academics, and human rights defenders).[50] Individuals targeted by these campaigns are often already vulnerable to discriminatory conditions, making the cost of public life too high. In addition to women and girls, targeted communities include advocates of so-called "gender ideology", who stand in stark opposition to groups that oppose LGBT+ rights, which are regarded as a critical site for the defence of liberal values.[51] These campaigns range from various tactics of using gendered stereotypes to vilify advocates to more specifically targeted techniques of spreading lies in conjunction with mobilizing hate speech and cyberbullying against individuals.[52] By reinforcing biases, such narratives targeting gender identities also perpetuate a political climate conducive to discrimination and hostility, which tends to make things worse for marginalized communities.

Kamala Harris, for instance, was subjected to large-scale social media smear campaigns that questioned her legitimacy and competence by attacking her racial identity, and her conformity to femininity and motherhood when she ran for vice president in 2020. Following her nomination, Harris's selection was denigrated through social media campaigns such as #HeelsUpHarris and #JoeandtheHoe, which spread sexualized and misogynistic narratives suggesting that her political career was the result of how she "slept her way to the top".[53] These narratives were found to be amplified by low-credibility social media accounts, including bots and trolls, aimed at discrediting her candidacy. These narratives, combined with new false stories, such as a fake hit-and-run incident that originated from Russian state media,[54] resurfaced during the 2024 presidential campaign to cast doubt on Harris's background, morality, and professionalism. In September 2024, the Microsoft Threat Analysis Center (MTAC) reported that Russian actors, allegedly linked to the Internet Research Agency, were intensifying their efforts to discredit Harris by amplifying compromising narratives and controversies surrounding her presidential campaign.[55] US intelligence sources subsequently reported attempted influence operations by Russia and Iran during the 2024 election, aimed at compromising the campaigns of

50 Elsa Hedling, 'Gendered Disinformation', *Feminist Foreign Policy Analysis: A New Subfield*, 2024, 137.

51 Cecilia Strand et al., *Disinformation Campaigns about LGBTI+ People in the EU and Foreign Influence* (European Parliament, Policy Department for External Relations, 2021).

52 European External Action Service, 'FIMI Targeting LGBTIQ+ People: Well-Informed Analysis to Protect Human Rights and Diversity', October 2023, https://www.eeas.europa.eu/eeas/fimi-targeting-lgbtiq-people_en.

53 Rebecca Coates Nee, '# JoeandtheHoe: Exploring Gender and Racial Stereotypes Used to Discredit Kamala Harris in the 2020 Presidential Election', *Howard Journal of Communications* 34, no. 3 (2023): 273–292; Nina Jankowicz et al., *Malign Creativity: How Gender, Sex, and Lies Are Weaponized against Women Online* (Woodrow Wilson International Center for Scholars, 2021).

54 The Office of the Director of National Intelligence, 'Election Security Update as of Mid-September 2024', September 2024, https://www.odni.gov/files/FMIC/documents/ODNI-Election-Security-Update-20240923.pdf.

55 Microsoft Threat Analysis Centre, 'Russia Leverages Cyber Proxies and Volga Flood Assets in Expansive Influence Efforts', September 2024, https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/msc/documents/presentations/CSR/MTAC-Election-Report-4.pdf.

both candidates.[56] However, these actors were able to exploit polarizing narratives about race and gender to target Harris's campaign in particular.

Other ways in which gender bias is mobilized include tactics that activate homophobia and transphobia to demean or humiliate individuals or state representatives. For example, French President Emmanuel Macron has been targeted by smear campaigns alleging that he was involved in an "extramarital gay relationship" and that his political campaign was backed by "the wealthy gay lobby". [57] During Canada's leadership of NATO's Enhanced Forward Presence mission in Latvia, Russian websites spread homophobic and transphobic narratives and reports, as well as photographs of convicted murderer and former Air Force Commander Russel Williams in women's underwear.[58] The aim was to spread fear that the Canadian troops were dominated by homosexuals and degenerates, and were not to be trusted by the Latvians.[59] This tactic was intended to resonate with existing homophobia and beliefs in Western moral decay to induce fear and distrust of NATO troops among the local population.

Disinformation that follows this pattern of activating gender biases goes beyond targeted attacks against women in visible public roles as part of a broader strategy to weaponize gender and identity politics. By effectively portraying women and gender non-conforming individuals as "unfit" for leadership or positions of trust, threat actors can activate deep-seated prejudices to increase division, spread mistrust, and deter marginalized groups or individuals from political participation.[60] Some of these campaigns also target broader movements for gender equality, sexual and reproductive health and rights, and LGBT+ rights, framing these issues as a geopolitical fault line.[61]

4.4 Muslim communities under threat in Sweden: Religious identity and political ideology

In 2021–2022, a disinformation campaign specifically targeting Swedish social services and the Care of Young Persons Act (Special Provisions) (LVU) spread rapidly in Sweden and abroad. The campaign's central narrative alleged that Swedish social services were systematically taking custody of children from immigrant families, especially Muslim

56 The Director of National Intelligence, 'Joint ODNI, FBI, and CISA Statement', November 2024, https://www.cisa.gov/news-events/news/joint-odni-fbi-and-cisa-statement-1.

57 Jean-Baptiste Jeangène Vilmer, 'The "Macron Leaks" Operation: A Post-Mortem' (Atlantic Council, 2019), 4–5, https://www.atlanticcouncil.org/wp-content/uploads/2019/06/The_Macron_Leaks_Operation-A_Post-Mortem.pdf.

58 Chris Brown, 'Anti-Canada Propaganda Greets Troops in Latvia', June 2017, https://www.cbc.ca/news/world/latvia-propaganda-1.4162612.

59 Owen Wong, 'Gender-Based Disinformation: A Tool of Hybrid Warfare' (Centre for International and Defence Policy at Queen's University, February 2024), https://www.queensu.ca/cidp/sites/cidpwww/files/uploaded_files/9–1%20CIPD%20Policy%20Brief%20Wong%20(Revised%20July%202024).pdf.

60 Hedling, 'Gendered Disinformation'.

61 Emil Edenborg, 'Disinformation and Gendered Boundarymaking: Nordic Media Audiences Making Sense of "Swedish Decline"', *Cooperation and Conflict* 57, no. 4 (2022): 496–515.

families, without any legal basis. The campaign exposed the vulnerabilities of mistrust of ethnic and religious minority groups in Sweden. The narrative was reinforced with emotionally charged warnings to the targeted groups that "Swedish public institutions kidnap children", which spread among Muslim communities in Sweden. As these communities mobilized to protest against the Swedish authorities through demonstrations and on social media, international actors amplified the outrage. Most notably, influencers with links to radical Islamist milieus, some of which were known to be violent, helped spread the campaign online.[62] The rapid international dissemination drew attention and engagement, especially from Muslim communities in the Middle East and North Africa. The swift spread and massive reach of the campaign led to it being described by the Swedish Psychological Defence Agency as one of the most serious cases of disinformation that Sweden had faced in a long time.[63]

This case demonstrates how disinformation campaigns can exploit the intersection of social identities and institutional distrust among communities to undermine and weaken democracies. By targeting religious and ethnic identities, this campaign mobilized and further fuelled fear, distrust, and alienation in immigrant and minority communities. The exploitation of minority identities, and the effective framing of Muslims as victims of systemic abuse in Sweden, greatly increased the emotional impact of the campaign both in Sweden and among Muslim communities worldwide. This framing of an immediate threat to Muslim families made the disinformation more damaging and difficult to contain, despite efforts to debunk it.

The situation was exacerbated by a series of Quran burnings in Sweden in 2023, which were reported by both national and international media. In January 2023, a Danish far-right politician burned a copy of the Quran in front of the Turkish embassy in Stockholm. The event took place during the process of Sweden's accession to NATO, leading Turkey's defence minister to call for the suspension of NATO membership negotiations with Sweden. In addition to stirring anger and fear among ethnic and religious groups in Sweden and abroad, the Quran burnings reignited an ideological debate over the conflict between individual and group rights, specifically free speech, religious freedom, and censorship. The Quran burnings and the ensuing debates consequently reproduced the tensions and divisions following the LVU campaign, which were exploited by new threat actors. Among a series of interference attempts, a cyberattack by a group linked to Iran's Revolutionary Guard hacked a bulk text messaging service and sent messages to approximately 15,000 people in Sweden in 2023.[64] The messages encouraged people to take revenge against

62 Magnus Ranstorp and Linda Ahlerup, 'The LVU Campaign: The Disinformation Campaign against Swedish Social Services: Disinformation, Conspiracy Theories and Domestic-International Links Related to Malign Influence Activities by Non-State Actors', (Försvarshögskolan, 2024).

63 Government Offices of Sweden, 'Government Taking Strong Action against Disinformation and Rumour-Spreading Campaign', February 2023, https://www.government.se/press-releases/2023/02/government-taking-strong-action-against-disinformation-and-rumour-spreading-campaign/.

64 Åklagarmyndigheten, Grovt dataintrång utfört av Iran (24 September 2024).

those who burn the Quran to further increase intergroup threat perceptions, and portrayed Sweden as an anti-Muslim country to undermine cohesion.[65]

This case demonstrates how the manipulation and exploitation of identity-based grievances, mistrust and fear can spiral into social and ideological division. In this case, the disinformation campaigns, and those who helped to amplify their effects, exploited fears of discrimination and portrayed the Swedish authorities as oppressive and hostile to specific communities. This exposed a vulnerability in Swedish society in terms of a lack of trust in the authorities and the welfare system among sectors of the population. By exploiting a vulnerability that was already there, the mistrust could be further reinforced in a way that risked undermining public institutions and radicalizing affected groups. This strategy of weaponizing identity is a tactic to destabilize democracies from within by weakening the effectiveness of democratic governance. The fact that these events included a stalled NATO accession process demonstrates the severity and potential security implications of exploiting democratic vulnerabilities.

65 Swedish Secruity Service, 'Foreign Power Influence Campaign Carried out through Breach of Data Security', September 2024, https://sakerhetspolisen.se/ovriga-sidor/other-languages/english-engelska/press-room/news/news/2024-09-25-foreign-power-influence-campaign-carried-out-through-breach-of-data-security.html.

5. Conclusions

This paper has highlighted the important role that social identities play in processes of pluralism and polarization in democracy. This role, and the diversity reflected in the representation of multiple and intersecting social identities in society, is both a strength and a vulnerability in the context of hybrid threats. A deeper understanding of this duality and how social identities are instrumental in both pluralism and polarization, especially in affective polarization, can provide important insights into how identity politics intersects with hybrid threats.

What can be learned from the examples highlighted in this paper is that the role of social identities in reinforcing or undermining democratic processes oscillates like a pendulum. Social categorization and polarization are not inherently negative, but their influence can be manipulated, especially in times of heightened emotions and uncertainty. Moreover, different types of crises can create conditions for systemic vulnerabilities across political systems and over time. This calls for a context-sensitive approach to analyzing when and how vulnerabilities emerge and how best to address them. An important step is to recognize and understand how these processes might shift and interact with factors such as strategic timing, levels of uncertainty and trust, and the "public mood".[66] These factors collectively influence how decision-makers can and should effectively deter and respond to hybrid threats. The real-life examples in this paper also highlight the importance of understanding intersectional dynamics and the complexity of how intergroup threat perceptions can evolve and reproduce along multiple politicized lines. These dynamics further complicate effective responses and strategies for containing interference attempts.

While the exploitation of identity politics by both domestic and international threat actors is not a new phenomenon or even restricted to the digital age, its manifestations and harmful effects have increased significantly due to the modern global and interconnected information environment. As many of the examples in this paper demonstrate, the speed and scale of the spread of misinformation and disinformation poses a challenge to the authorities and organizations tasked with monitoring and countering its influence. In addition, the weaponization of identity politics exposes inequalities, biases and social grievances that reveal democratic shortcomings. To effectively respond to, mitigate and deter hybrid threats, states must also engage in introspective analysis to better understand and address vulnerabilities. Strategies to address democratic vulnerabilities should therefore also promote long-term cohesion by strengthening democratic governance systems and building inter-societal trust between communities and democratic institutions. Increasing levels of trust is crucial for reducing democratic vulnerabilities and strengthening the shield of "pluralistic resilience" in times of turbulence.

66 Hedling and Ördén, 'Disinformation, Deterrence and the Politics of Attribution'.

Author

Elsa Hedling is an Associate Professor of Political Science and an Associate Senior Lecturer in European Studies at the Centre for Languages and Literature, Lund University. She is also affiliated with the Psychological Defence Research Institute (PDRI) and the Swedish Institute of International Affairs in Stockholm. Her current research focuses on democratic resilience, hybrid threats, and the role of diplomacy and multilateral cooperation in a rapidly digitalizing world. She has published articles on these topics in leading academic journals such as *International Affairs* (2021 and 2025), *Review of International Studies* (2022), *International Studies Quarterly* (2024), and *Global Studies Quarterly* (2022 and 2024). Elsa has also co-authored two books: one on EU foreign and security policy (Edward Elgar, 2022) and the other on feminist foreign policy and digital diplomacy (Palgrave Macmillan, 2024).

Hybrid CoE

The European Centre of Excellence
for Countering Hybrid Threats